



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 192050

TO: Ruixiang Li
Location: REM/4D75/4C70
Art Unit: 1646
Monday, June 12, 2006
Case Serial Number: 10/600816

From: Dwayne Hill
Location: Biotech-Chem Library
REM-1A64
Phone: (571) 272-8972

Dwayne.hill@uspto.gov

Search Notes

Examiner Li,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Dwayne Hill
Technical Information Specialist
STIC Biotech/Chem Library
(571) 272-8972

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

199050

From: Li, Ruixiang
Sent: Monday, June 05, 2006 3:22 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search of Application No.10/600,816

Please do a standard search on:

- (i). SEQ ID NO: 3 against amino acid databases;
- (ii). SEQ ID NOS: 2 and 30-33 against nucleic acid databases.

Thank you very much!

Ruixiang Li
GAU 1646
REM 4D75
Mail Box 4C70
(571) 272-0875

STIC-BIOTECH/CHEMLIB
JUN 5 2006
10:00 AM

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

QY 2359 TGGAGAGTCTTCATTGTATAAAGTATTACCTGAGTATGCAATAAGATGTGGTG 2418
|||||
Db 24001 TGGAGAGTCTTCATTGTATAAAGTATTACCTGAGTATGCAATAAGATGTGGTG 24060
|||||
QY 2419 CCACTCTTTTCATGGTGGGCGACA 2443
|||||
Db 24061 CCACTCTTTTCATGGTGGGCGACA 24085
|||||

RESULT 4

US-09-949-016-12314
; Sequence 12314, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12314
; LENGTH: 26238
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12314

Query Match 49.5%; Score 1215.4; DB 3; Length 26238;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1219; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1219 ACATTTTACGTGCGAGAACAGGCTCCCAAAAGGAATTTCTCATCCCGGCGCCACGC 1278
|||||
Db 23015 ACTGTTTCTTTTTCAGAACAGGCTCCCAAAAGGAATTTCTCATCCCGGCGCCACGC 23074
|||||
QY 1279 TTGGCGAGCCCTTACAAAGACTATCAAGTAAAGAGGGGAGGAGTCTGCTCTGA 1338
|||||
Db 23075 TTGGCGAGCCCTTACAAAGACTATCAAGTAAAGAGGGGAGGAGTCTGCTCTGA 23134
|||||
QY 1339 AGAGTGGGACAAATGCGCGCGCGCAGATCTAGCGGGAGCTCAAGGGAGTGGGCGA 1398
|||||
Db 23135 AGAGTGGGACAAATGCGCGCGCGCAGATCTAGCGGGAGCTCAAGGGAGTGGGCGA 23194
|||||
QY 1399 AATCTTGAAGTCTTCTGAGAAAACGTGTACAGACACTACGGGAACAGTTTGCCTCCCTCC 1458
|||||
Db 23195 AATCTTGAAGTCTTCTGAGAAAACGTGTACAGACACTACGGGAACAGTTTGCCTCCCTCC 23254
|||||
QY 1459 AGCCTCAACCAAAATTTCTTCATGCTGGGCTGATGCTGGCTAGTAAAGTCCAGTTCTT 1518
|||||
Db 23255 AGCCTCAACCAAAATTTCTTCATGCTGGGCTGATGCTGGCTAGTAAAGTCCAGTTCTT 23314
|||||
QY 1519 AGAGGCGCTGTAGTATTTTCTTCTCATCTCTTGGATCTCTTTTAAAGTGGG 1578
|||||
Db 23315 AGAGGCGCTGTAGTATTTTCTTCTCATCTCTTGGATCTCTTTTAAAGTGGG 23374
|||||
QY 1579 AGTCTCAGGCAACTCAAGTTTACAGCCCTTACTCTTTTGTGTTTGTGTTTGAACAGGATC 1638
|||||
Db 23375 AGTCTCAGGCAACTCAAGTTTACAGCCCTTACTCTTTTGTGTTTGTGTTTGAACAGGATC 23434
|||||
QY 1639 TTGCTCTGTCACCCAGGCTTGAGTCAGTGGTGGATCAGCCCGCAGTGCAGCTCGAC 1698
|||||
Db 23435 TTGCTCTGTCACCCAGGCTTGAGTCAGTGGTGGATCAGCCCGCAGTGCAGCTCGAC 23494
|||||
QY 1699 ACCTGTGCTCAAGCAATCTCCATCTCCATCTCCCAAGTGTGGGATGACAGGCGTGA 1758
|||||

Db 23495 ACCTGTGCTCAAGCAATCTCCATCTCCATCTCCCAAGTGTGGATGACAGGCGTGA 23554
QY 1759 GCCACAGCTCCACAGCCTAGGCCCTTAATCTTGTCTGTATTATTTTCATGGAATAAGTCTG 1818
|||||
Db 23555 GCCACAGCTCCACAGCCTAGGCCCTTAATCTTGTCTGTATTATTTTCATGGAATAAGTCTG 23614
|||||
QY 1819 GTCATCTGAGCTCACGCTGGCTCACACAGCTCTAGGGGCTGCTCTCTAACTCACAGTG 1878
|||||
Db 23615 GTCATCTGAGCTCACGCTGGCTCACACAGCTCTAGGGGCTGCTCTCTAACTCACAGTG 23674
|||||
QY 1879 GGTTTTGTGAGGCTCTGTGGCCCGAGAGACCTGCAATCTGAGCAAAAATAAGCAAAAG 1938
|||||
Db 23675 GGTTTTGTGAGGCTCTGTGGCCCGAGAGACCTGCAATCTGAGCAAAAATAAGCAAAAG 23734
|||||
QY 1939 CTTCTCTAGCCGACTGGCTGAACTACACTGGAAGCCAACTTGTGTGGACCCCGCTC 1998
|||||
Db 23735 CTTCTCTAGCCGACTGGCTGAACTACACTGGAAGCCAACTTGTGTGGACCCCGCTC 23794
|||||
QY 1999 CCCAACCCCTTTTGGCTGGTAGGAGGCTAAAGATCACCCCTAAATTTTACTCATCTCTC 2058
|||||
Db 23795 CCCAACCCCTTTTGGCTGGTAGGAGGCTAAAGATCACCCCTAAATTTTACTCATCTCTC 23854
|||||
QY 2059 TAGTGTGCTCTCACATTTGGGCTCAGCAGCTCCCGAGCACCATTTCACAGGTCAACCCCTC 2118
|||||
Db 23855 TAGTGTGCTCTCACATTTGGGCTCAGCAGCTCCCGAGCACCATTTCACAGGTCAACCCCTC 23914
|||||
QY 2119 TCTTCTTGACACTGTCCCAAACTTGTGTCAATTCGAGATCTAACTCTCCCTCTAGCCTC 2178
|||||
Db 23915 TCTTCTTGACACTGTCCCAAACTTGTGTCAATTCGAGATCTAACTCTCCCTCTAGCCTC 23974
|||||
QY 2179 TGCAGGAATTTCTTTCAGACCTCAGTAGCACAAAGCCGGTTGCTCTTGTGAGGAGATT 2238
|||||
Db 23975 TGCAGGAATTTCTTTCAGACCTCAGTAGCACAAAGCCGGTTGCTCTTGTGAGGAGATT 24034
|||||
QY 2239 TGTAGATCAATCTCACTTCAAAATTCCTGGGCTGATATCTCTCATCTTGCACCCCAAC 2298
|||||
Db 24035 TGTAGATCAATCTCACTTCAAAATTCCTGGGCTGATATCTCTCATCTTGCACCCCAAC 24094
|||||
QY 2299 CTCTGTAATAGATTACCGCATTTACGGCTGCAATTTCTGTAAGTGGCATGGTCTCTAA 2358
|||||
Db 24095 CTCTGTAATAGATTACCGCATTTACGGCTGCAATTTCTGTAAGTGGCATGGTCTCTAA 24154
|||||
QY 2359 TGGAGAGTGTTCATTGTATAAAGTATTACCTGAGTATGCAATAAAGATGTGGTG 2418
|||||
Db 24155 TGGAGAGTGTTCATTGTATAAAGTATTACCTGAGTATGCAATAAAGATGTGGTG 24214
|||||
QY 2419 CCACTCTTTTCATGGTGGGCGACA 2443
|||||
Db 24215 CCACTCTTTTCATGGTGGGCGACA 24239
|||||

RESULT 5

US-09-188-930-249
; Sequence 249, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.101c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-188-930-249

Query Match		45.8%;	Score 1126;	DB 3;	Length 1212;	
Best Local Similarity		99.1%;	Pred. No. 2.8e-305;			
Matches 1132;		Conservative	0;	Mismatches 10;	Indels	0; Gaps 0;
QY	297	CCAAGTACTACAGACTTTGTGTAAGGCTGAAGCTTGGGGCATCGTCCCTAGAAAAAGGTGG	356			
DB	16	CGAGGTACTACAGACTTTGTGTAAGGCTGAAGCTTGGGGCATCGTCCCTAGAAAAAGGTGG	75			
QY	357	CCACAGCGGGGTGTGACCTCGGTGGCTTCATGCTCACTCCCGATCCTCGTCTGCA	416			
DB	76	CCACAGCTGGGTGTGACCTCGGTGGCTTCATGCTCACTCCCGATCCTCGTCTGCA	135			
QY	417	AGGTGAGGACTCCAAAGCGGAAAAATGCTGCCTACTCAGTTCCTCTTCTCTCTGGGTG	476			
DB	136	AGGTGAGGACTCCAAAGCGGAAAAATGCTGCCTACTCAGTTCCTCTTCTCTCTGGGTG	195			
QY	477	TGTTGGGCATCTTGGCCTCACCTTCGCTTCATCATCGGACTGGACGGGAGCACAGGC	536			
DB	196	TGTTGGGCATCTTGGCCTCACCTTCGCTTCATCATCGGACTGGACGGGAGCACAGGC	255			
QY	537	CCACAGCTTCTTCCCTTTGGATCCTCTTTTCCATCTGCTTCTCCTCGCTGGCTC	596			
DB	256	CCACAGCTTCTTCCCTTTGGATCCTCTTTTCCATCTGCTTCTCCTCGCTGGCTC	315			
QY	597	ATGCTGTCACTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGATTC	656			
DB	316	ATGCTGTCACTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGATTC	375			
QY	657	TGGGTCTGGCGTGGCTTACGCTAGTCCAGATGTTATCGCTATTGAATATATTGTC	716			
DB	376	TGGGTCTGGCGTGGCTTACGCTAGTCCAGATGTTATCGCTATTGAATATATTGTC	435			
QY	717	TGACCATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCCGCTCCTCGTGAATG	776			
DB	436	TGACCATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCCGCTCCTCGTGAATG	495			
QY	777	AAGACTTTGCTCCTGCTCAGCTACGCTCCTCTTTGTAAGGCTGACCTTCTCCTCATG	836			
DB	496	AAGACTTTGCTCCTGCTCAGCTACGCTCCTCTTTGTAAGGCTGACCTTCTCCTCATG	555			
QY	837	CCTCTTCACTTCTGTGTTCTTACGGGCTGGAGAGACATGGGGCCACATCTACC	896			
DB	556	CCTCTTCACTTCTGTGTTCTTACGGGCTGGAGAGACATGGGGCCACATCTACC	615			
QY	897	TCAGATGCTCCTCTCAATTGGCATCTGGGTGGCTTGGATCACCTCTCATGCTTCTCTG	956			
DB	616	TCAGATGCTCCTCTCAATTGGCATCTGGGTGGCTTGGATCACCTCTCATGCTTCTCTG	675			
QY	957	ACTTTGACCGGAGTGGGATGACACATCCTCAGCTCCGCTTGGGTGCAATGGCTGGG	1016			
DB	676	ACTTTGACCGGAGTGGGATGACACATCCTCAGCTCCGCTTGGGTGCAATGGCTGGG	735			
QY	1017	TGTTCTGTGTTGCTTATGTTAGTCCGAGTTTGGCTGCTCAAGAGCAACGAAACCCCA	1076			
DB	736	TGTTCTGTGTTGCTTATGTTAGTCCGAGTTTGGCTGCTCAAGAGCAACGAAACCCCA	795			
QY	1077	TGGAATTATCCTGTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTG	1136			
DB	796	TGGAATTATCCTGTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTG	855			
QY	1137	TGGAAGACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGAGAGGGGACA	1196			
DB	856	TGGAAGACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGAGAGGGGACA	915			
QY	1197	CGCTCTATGCCCTTATTCACACATTTTCAGTGCAGAACAGCCTCCCTCCCAAGGAAT	1256			
DB	916	CGCTCTATGCCCTTATTCACACATTTTCAGTGCAGAACAGCCTCCCTCCCAAGGAAT	975			
QY	1257	TCTCCATCCACGGGCCCCAGCTTGGCGAGGCCCTTTACAAAGACTATGAAGTAAAGAAAG	1316			
DB	976	TCTCCATCCACGGGCCCCAGCTTGGCGAGGCCCTTTACAAAGACTATGAAGTAAAGAAAG	1035			

QY	1317	AGGGCAGCTAACTCTGTCTCTGAAGAGTGGGACAAATGACGCCGGGGCGGAGATCTAGCGG	1376			
DB	1036	AGGGCAGCTAACTCTGTCTCTGAAGAGTGGGACAAATGACGCCGGGGCGGAGATCTAGCGG	1095			
QY	1377	GAGCTCAAAGGATGTGGCGGAAATCTTCTGAGTCTTCTGAGAAACTGTACAACACACTAC	1436			
DB	1096	GAGCTCAAAGGATGTGGCGGAAATCTTCTGAGTCTTCTGAGAAACTGTACCTCCCGGGC	1155			
QY	1437	GG 1438				
DB	1156	GG 1157				
RESULT 6						
US-09-312-283C-249						
; Sequence 249, Application US/09312283C						
; Patent No. 6573095						
; GENERAL INFORMATION:						
; APPLICANT: Watson, James D.						
; APPLICANT: Strachan, Lorna						
; APPLICANT: Sleeman, Matthew						
; APPLICANT: Onrust, Rene						
; APPLICANT: Murison, James G.						
; APPLICANT: Kumble, Krishanand D.						
; TITLE OF INVENTION: Compositions Isolated from Skin Cells						
; FILE REFERENCE: 11000.1011c2						
; CURRENT APPLICATION NUMBER: US/09/312,283C						
; NUMBER OF SEQ ID NOS: 425						
; SOFTWARE: FastaSeq for Windows Version 4.0						
; SEQ ID NO 249						
; LENGTH: 1212						
; TYPE: DNA						
; ORGANISM: Mouse						
US-09-312-283C-249						
Query Match						
Best Local Similarity 99.1%;						
Matches 1132; Conservative						
0; Mismatches 10; Indels 0; Gaps 0;						
QY	297	CCAAGTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCCCTAGAAAAAGGTGG	356			
DB	16	CGAGGTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCCCTAGAAAAAGGTGG	75			
QY	357	CCACAGCGGGGTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCA	416			
DB	76	CCACAGCTGGGTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCA	135			
QY	417	AGGTGAGGACTCCAAAGCGGAAAAATGCTGCTCACTCTCCGATCCTCGTCTGCA	476			
DB	136	AGGTGAGGACTCCAAAGCGGAAAAATGCTGCTCACTCTCCGATCCTCGTCTGCA	195			
QY	477	TGTTGGGCATCTTGGCCTCACCTTCGCTTCATCATCGGACTGGACGGGAGCACAGGC	536			
DB	196	TGTTGGGCATCTTGGCCTCACCTTCGCTTCATCATCGGACTGGACGGGAGCACAGGC	255			
QY	537	CCACAGCTTCTTCCCTTTGGGATCCTCTTTTCCATCTGCTTCTCCTCGCTGGCTC	596			
DB	256	CCACAGCTTCTTCCCTTTGGGATCCTCTTTTCCATCTGCTTCTCCTCGCTGGCTC	315			
QY	597	ATGCTGTCACTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGATTC	656			
DB	316	ATGCTGTCACTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGATTC	375			
QY	657	TGGGTCTGGCGTGGCTTACGCTAGTCCAGATGTTATCGCTATTGAATATATTGTC	716			
DB	376	TGGGTCTGGCGTGGCTTACGCTAGTCCAGATGTTATCGCTATTGAATATATTGTC	435			
QY	717	TGACCATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCCGCTCCTCGTGAATG	776			
DB	436	TGACCATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCCGCTCCTCGTGAATG	495			
QY	777	AAGACTTTGCTCCTGCTCAGCTACGCTCCTCTTTGTAAGGCTGACCTTCTCCTCATG	836			
DB	496	AAGACTTTGCTCCTGCTCAGCTACGCTCCTCTTTGTAAGGCTGACCTTCTCCTCATG	555			
QY	837	CCTCTTCACTTCTGTGTTCTTACGGGCTGGAGAGACATGGGGCCACATCTACC	896			
DB	556	CCTCTTCACTTCTGTGTTCTTACGGGCTGGAGAGACATGGGGCCACATCTACC	615			
QY	897	TCAGATGCTCCTCTCAATTGGCATCTGGGTGGCTTGGATCACCTCTCATGCTTCTCTG	956			
DB	616	TCAGATGCTCCTCTCAATTGGCATCTGGGTGGCTTGGATCACCTCTCATGCTTCTCTG	675			
QY	957	ACTTTGACCGGAGTGGGATGACACATCCTCAGCTCCGCTTGGGTGCAATGGCTGGG	1016			
DB	676	ACTTTGACCGGAGTGGGATGACACATCCTCAGCTCCGCTTGGGTGCAATGGCTGGG	735			
QY	1017	TGTTCTGTGTTGCTTATGTTAGTCCGAGTTTGGCTGCTCAAGAGCAACGAAACCCCA	1076			
DB	736	TGTTCTGTGTTGCTTATGTTAGTCCGAGTTTGGCTGCTCAAGAGCAACGAAACCCCA	795			
QY	1077	TGGAATTATCCTGTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTG	1136			
DB	796	TGGAATTATCCTGTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTG	855			
QY	1137	TGGAAGACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGAGAGGGGACA	1196			
DB	856	TGGAAGACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGAGAGGGGACA	915			
QY	1197	CGCTCTATGCCCTTATTCACACATTTTCAGTGCAGAACAGCCTCCCTCCCAAGGAAT	1256			
DB	916	CGCTCTATGCCCTTATTCACACATTTTCAGTGCAGAACAGCCTCCCTCCCAAGGAAT	975			
QY	1257	TCTCCATCCACGGGCCCCAGCTTGGCGAGGCCCTTTACAAAGACTATGAAGTAAAGAAAG	1316			
DB	976	TCTCCATCCACGGGCCCCAGCTTGGCGAGGCCCTTTACAAAGACTATGAAGTAAAGAAAG	1035			

777 AAGACTTTGTCTCCTGCTCAGCTACGTCCTCTTCTTGATGGCGCTGACCTCTCCTCATGT 836
|||||
496 AAGACTTTGTCTCCTGCTCAGCTACGTCCTCTTCTTGATGGCGCTGACCTCTCCTCATGT 555
|||||
837 CTTCTTTACCTTTCTGTGGTTCCTTCAOGGCTTGAAGAGACATGGGGCCCACTTACC 896
|||||
556 CTTCTTTACCTTTCTGTGGTTCCTTCAOGGCTTGAAGAGACATGGGGCCCACTTACC 615
|||||
897 TCAGGATGCTCCTCTCCATGTCCTGCTGGGTGGCTGGATCACCTGCTCATGCTTCCTG 956
|||||
616 TCAGGATGCTCCTCTCCATGTCCTGCTGGGTGGCTGGATCACCTGCTCATGCTTCCTG 675
|||||
957 ACTTTGACCCAGCTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGG 1016
|||||
676 ACTTTGACCCAGCTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGG 735
|||||
1017 TGTCTCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAAAAGCAACGAAACCCCA 1076
|||||
736 TGTCTCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAAAAGCAACGAAACCCCA 795
|||||
1077 TGGATTATCCTGTTGAGGATGCTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTG 1136
|||||
796 TGGATTATCCTGTTGAGGATGCTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTG 855
|||||
1137 TGGAGAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGGACA 1196
|||||
856 TGGAGAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGGACA 915
|||||
1197 CGCTTATGCCCCCTATTCCACACATTTTCAGCTGCAGAACCCAGCTTCCCCAAAGGAAT 1256
|||||
916 CGCTTATGCCCCCTATTCCACACATTTTCAGCTGCAGAACCCAGCTTCCCCAAAGGAAT 975
|||||
1257 TCTCCATCCACGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAG 1316
|||||
976 TCTCCATCCACGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAG 1035
|||||
1317 AGGGCAGCTAACTGTCTGTAAGAGTGGGACAAATCGACCCGGCGGAGATCTAGCGG 1376
|||||
1036 AGGGCAGCTAACTGTCTGTAAGAGTGGGACAAATCGACCCGGCGGAGATCTAGCGG 1095
|||||
1377 GAGCTCAAAAGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTAC 1436
|||||
1096 GAGCTCAAAAGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACTCGCCCGGC 1155
|||||
1437 GG 1438
1156 GG 1157

RESULT 7
US-09-495-050A-223
; Sequence 223, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice.
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6492505 2681738CT1
US-09-495-050A-223

Query Match 42.0%; Score 1031.6; DB 3; Length 1228;
Best Local Similarity 98.6%; Pred. No. 8.3e-279;
Matches 1094; Conservative 0; Mismatches 9; Indels 7; Gaps 5;
QY 732 CCNAACGTCATATGCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTTGTCTCC 791
DB 1 CCAACGTCATATGCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTTGTCTCC 60
QY 792 TGTCTACCTTACGTCCTTCTTCTGATGGCGTGACCTTCTCTCATGTCTCTCTTCACTTCT 851
DB 61 TGTCTACCTTACGTCCTTCTTCTGATGGCGTGACCTTCTCTCATGTCTCTCTTCACTTCT 120
QY 852 GTGGTTCCTTACGGGCTGGAAGACATGCGGGCCACATCTACCTACCGATGCTCTCT 911
DB 121 GTGGTTCCTTACGGGCTGGAAGACATGCGGGCCACATCTACCTACCGATGCTCTCT 180
QY 912 CCATTGCCATCTGGGTGGCTGGATCACCTCGTCTCATGTCTTCTGATCTTTGACCGCAGGT 971
DB 181 CCATTGCCATCTGGGTGGCTGGATCACCTCGTCTCATGTCTTCTGATCTTTGACCGCAGGT 240
QY 972 G-GGATGACACCATCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGGTGTTCTCTTGGCT 1030
DB 241 GAGGATGACACCATCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGGTGTTCTCTTGGCT 300
QY 1031 TATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGATTTATCTCTTT 1090
DB 301 TATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGATTTATCTCTTT 360
QY 1091 GAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGCTGTGGAGAACAGAGCC 1150
DB 361 GAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGCTGTGGAGAACAGAGCC 420
QY 1151 TACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGGACAGCTCTATATGCCCC 1210
DB 421 TACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGGACAGCTCTATATGCCCC 480
QY 1211 TATTTCCACATTTTTCAGCTGCGAGAACCCAGCTTCCCCAAAGGAAATTTCTCCATCCACGG 1270
DB 481 TATTTCCACATTTTTCAGCTGCGAGAACCCAGCTTCCCCAAAGGAAATTTCTCCATCCACGG 540
QY 1271 GCCCAGCTTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGCGAGCTAACTC 1330
DB 541 GCCCAGCTTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGCGAGCTAACTC 600
QY 1331 TGTCTTGAAGAGTGGGACAAATCGACCGGGCGGCGAGATCTAGCGGAGCTCAAGGGAT 1390
DB 601 TGTCTTGAAGAGTGGGACAAATCGACCGGGCGGCGAGATCTAGCGGAGCTCAAGGGAT 660
QY 1391 GTGGGCGAAATCTTGAGTCTTCTGAGAAACTCTGTAAGACACTACCGGGAACAGTTTGGCC 1450
DB 661 GTGGGCGAAATCTTGAGTCTTCTGAGAAACTCTGTAAGAACTGTAAGAACTGTTTGGCC 720
QY 1451 TCCTCTCCAGCTTCAACCAATTTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAAGACTC 1510
DB 721 TCCTCTCCAGCTTCAACCAATTTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAAGACTC 780
QY 1511 CAGTTCTTAGAGCGCTGTAGTA - TTTTCTTTTCTGTCATCTCTCTTGGATPACTTCT 1568
DB 781 CAGTTCTTAGAGCGCTGTAGTA - TTTTCTTTTCTGTCATCTCTCTTGGATPACTTCT 840
QY 1569 TTTTAAGTGGGAGTCTCAGGCACTCAAGTTTACACCTTACTCTTTTGTGTTTTTGTG 1628
DB 841 TTTTAAGTGGGAGTCTCAGGCACTCAAGTTTACACCTTACTCTTTTGTGTTTTTGTG 900
QY 1629 AAAACAGGATCTTGTCTGTGTCACCCAGGCTTGAAGTGGTGGATGCGATCACAGCCAGTGC 1688
DB 901 AAAACAGGATCTTGTCTGTGTCACCCAGGCTTGAAGTGGTGGATGCGATCACAGCCAGTGC 960
QY 1689 AGCTTCGACCATCTGTGCTCAAGCAATCTCTCCATCTCCCAATGCTGGGATG 1748
DB 961 AGCTTCGACCATCTGTGCTCAAGCAATCTCTCCATCTCTCCATCTCTCCCAATGCTGGGATG 1020

1749 ACAGGCGTAGCCACAGCTCCAGCCTAGGCCCTTAACTCTGCTCTTATTTTCCATGGAC 1808
 |||||
 1021 ACAGGCGTGA-CGACAGCTCCAGCCTAGGCCCTTAACTCTGCTTATTT--CCATGGC 1077
 |||||

1809 TAAAGGCTGTGCTCATCTGAGCTCAGCTGG 1838
 |||||

1078 CTAAAGGCTGTGTCA-CTGAGCTCAGCTGG 1106
 |||||

```

RESULT 8
US-09-949-016-29173
; Sequence 29173, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29173
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29173

```

Query Match	24.5%	Score 600.6	DB 3	Length 601
Best Local Similarity	99.8%	Pred. No. 3.9e-158		
Matches 600	Conservative 1	Mismatches 0	Indels 0	Gaps 0
QY	1258	CTCCATCCCACGGGCCACCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGA	1317	
DB	1	CTCCATCCCACGGGCCACCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGA	60	
QY	1318	GGGCGAGCTAACTCTGTCTCTGAAGAGTGGGCAAAATGCAGCGGGCGCGAGACTCTAGCGGG	1377	
DB	61	GGGCGAGCTAACTCTGTCTCTGAAGAGTGGGCAAAATGCAGCGGGCGCGAGACTCTAGCGGG	120	
QY	1378	AGCTCAAAAGGGATGTGGCGCAAACTCTTGAGTCTCTCAGAAAACTGTACAAAGACACTACG	1437	
DB	121	AGCTCAAAAGGGATGTGGCGCAAACTCTTGAGTCTCTCAGAAAACTGTACAAAGACACTACG	180	
QY	1438	GGAAACAGTTTGGCTCCCTCCAGCGCTCAACCAACAATCTCTTCATGCTGGGGCTGATGTGG	1497	
DB	181	GGAAACAGTTTGGCTCCCTCCAGCGCTCAACCAACAATCTCTTCATGCTGGGGCTGATGTGG	240	
QY	1498	GCTAGTAAGACTCCAGTCTTTAGAGGCGCTGTAGTATTTTTTTTTTTTTTGTCTCATCCTT	1557	
DB	241	GCTAGTAAGACTCCAGTCTTTAGAGGCGCTGTAGTATTTTTTTTTTTTTTGTCTCATCCTT	300	
QY	1558	TGATATCTCTTTTAAAGTGGGAGTCTCAGGGAACCTAAGTTTAGACCTTACTCTTTTTTG	1617	
DB	301	WGGATATCTCTTTTAAAGTGGGAGTCTCAGGGAACCTAAGTTTAGACCTTACTCTTTTTTG	360	
QY	1618	TTTGTGTTTTTGAACACAGATCTTGTCTGTGTCACCCAGGCTTGAGTGAGTGTCGATCA	1677	
DB	361	TTTGTGTTTTTGAACACAGATCTTGTCTGTGTCACCCAGGCTTGAGTGAGTGTCGATCA	420	
QY	1678	CAGCCCAAGTGCAGCCTCGACCACTGTGTGCTCAAGCAATCTCTCCCATCTCCATCTCCCAA	1737	
DB	421	CAGCCCAAGTGCAGCCTCGACCACTGTGTGCTCAAGCAATCTCTCCCATCTCCCAA	480	
QY	1738	GTGCTGGGATGACAGGGCGTGAGCCACAGCTCCAGCTCCAGCCCTTAATCTTGCTGTAT	1797	
DB	481	GTGCTGGGATGACAGGGCGTGAGCCACAGCTCCAGCTCCAGCCCTTAATCTTGCTGTAT	540	

Qy 1798 TTTCCATGGACTAAAGGCTCGGTCACTGAGCTCACGCTGGCTCACACAGCTCTAGGGGC 1857
 Db 541 TTTCCATGGACTAAAGGCTCGGTCACTGAGCTCACGCTGGCTCACACAGCTCTAGGGGC 600
 Qy 1858 C 1858
 Db 601 C 601

```

RESULT 9
US-09-949-016-29174
; Sequence 29174, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29174
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29174

```

Query Match	24.5%	Score 600.6	DB 3	Length 601
Best Local Similarity	99.8%	Pred. No. 3.9e-158		
Matches 600	Conservative 1	Mismatches 0	Indels 0	Gaps 0
Qy	1550	TCATCCCTTTGGATACCTCTCTTTAAAGTGGGAGCTCAGGCAACTCAAAGTTTAGACCTTAC	1609	
Db	1	TCATCCCTTTGGATACCTCTTTAAAGTGGGAGCTCAGGCAACTCAAAGTTTAGACCTTAC	60	
Qy	1610	TCCTTTTGTGTTTTTTTGAACAAGGATCTTGCTCTGTCAACCAGGCTTGAGTGAGTGG	1669	
Db	61	TCCTTTTGTGTTTTTTTGAACAAGGATCTTGCTCTGTCAACCAGGCTTGAGTGAGTGG	120	
Qy	1670	TGGATCACAGCCAGTCAGCCTTCGACACTGTGTCTCAAGCAATCCTCCCATCTCCAT	1729	
Db	121	TGGATCACAGCCAGTCAGCCTTCGACACTGTGTCTCAAGCAATCCTCCCATCTCCAT	180	
Qy	1730	CTCCAAAGTCTGGGATGACAGCGGTGAGCCA CAGCTCC CAGCCTAGGCCCTTAATCTT	1789	
Db	181	CTCCAAAGTCTGGGATGACAGCGGTGAGCCA CAGCTCC CAGCCTAGGCCCTTAATCTT	240	
Qy	1790	GCTGTTATTTTCCATGACATAAAGGTCTGGTCATCTGAGCTCA CGCTGGCTCACACAGCT	1849	
Db	241	GCTGTTATTTTCCATGACATAAAGGTCTGGTCATCTGAGCTCA CGCTGGCTCACACAGCT	300	
Qy	1850	CTAGGGGCGCTGCTCCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGGCCACAGAGCAGA	1909	
Db	301	YTAGGGGCGCTGCTCCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGGCCACAGAGCAGA	360	
Qy	1910	CTTGCAATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGGCCATGAGCTGAATCTACAC	1969	
Db	361	CCTGCATATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGGCCATGAGCTGAATCTACAC	420	
Qy	1970	TGGAAGCCAACTTCTGCGACCCCGGCTCCCCAACCCCTTCTTGCTGGGTAGGAGAGGCT	2029	
Db	421	TGGAAGCCAACTTCTGCGACCCCGGCTCCCCAACCCCTTCTTGCTGGGTAGGAGAGGCT	480	
Qy	2030	AAAGATCACCCCTAAATTTACTCTATCTCTAGTGTGCTCACATATGGGCCCTCAGAGCT	2089	

Db 481 AAAGATCACCCCTAAATTTACTCATCTCTAGTGTGCTCCCTCAATTTGGGCTCAGCAGCT 540
QY 2090 CCCAGACCAATTCAGGTACCCCTCTCTCTTCTTGCACTGTCCCAAACTTGTGTCA 2149
Db 541 CCCAGACCAATTCAGGTACCCCTCTCTCTTCTTGCACTGTCCCAAACTTGTGTCA 600
QY 2150 A 2150
Db 601 A 601

RESULT 10
US-09-949-016-126957
; Sequence 126957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 126957
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126957

Query Match 24.5%; Score 600.6; DB 3; Length 601;
Best Local Similarity 99.8%; Pred. No. 3.9e-158;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1258 CTCATCCACGGGCCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGA 1317
Db 1 CTCATCCACGGGCCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGA 60

QY 1318 GGGCAGCTAACTCTGTCTGAAGAGTGGGCAAAATCAGCGGGCGGCGAGATCAGCGGG 1377
Db 61 GGGCAGCTAACTCTGTCTGAAGAGTGGGCAAAATCAGCGGGCGGCGAGATCAGCGGG 120

QY 1378 AGCTCAAGGGATGTGGCGGAAATCTTGAGTCTTCTGAGAAAATCTGACAGACACTACG 1437
Db 121 AGCTCAAGGGATGTGGCGGAAATCTTGAGTCTTCTGAGAAAATCTGACAGACACTACG 180

QY 1438 GGAACAGTTTGCTCCCTCCAGCTCAACCAAAATCTTCCATGCTGGGGCTGATGTGG 1497
Db 181 GGAACAGTTTGCTCCCTCCAGCTCAACCAAAATCTTCCATGCTGGGGCTGATGTGG 240

QY 1498 GCTAGTAAGACTCCAGTTCTTAGAGGGCTGTAGTATTTTTTTTTTGTCTCATCCTT 1557
Db 241 GCTAGTAAGACTCCAGTTCTTAGAGGGCTGTAGTATTTTTTTTTTGTCTCATCCTT 300

QY 1558 TGGTACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTGG 1617
Db 301 WGGTACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTGG 360

QY 1618 TTTGTTTTTGAACAGGATCTTGTCTCTGTACCCAGGCTTGAGTGCAGTGGTGCATCA 1677
Db 361 TTTGTTTTTGAACAGGATCTTGTCTCTGTACCCAGGCTTGAGTGCAGTGGTGCATCA 420

QY 1678 CAGCCAGTGCAGCTCGACCCACTGTCTCAGCAATCTCTCCATCTCCATCTCCAAA 1737
Db 421 CAGCCAGTGCAGCTCGACCCACTGTCTCAGCAATCTCTCCATCTCCATCTCCAAA 480

QY 1738 GTGCTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTGTGTAT 1797

Db 481 GTGCTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTTAATCTTGTGTAT 540
QY 1798 TTTCCATGAGCTAAAGGTCTGTGTCATCTAGCTCAGCTGGCTCACACAGCTCTAGGGC 1857
Db 541 TTTCCATGAGCTAAAGGTCTGTGTCATCTAGCTCAGCTGGCTCACACAGCTCTAGGGC 600
QY 1858 C 1858
Db 601 C 601

RESULT 11
US-09-949-016-126958
; Sequence 126958, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 126958
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126958

Query Match 24.5%; Score 600.6; DB 3; Length 601;
Best Local Similarity 99.8%; Pred. No. 3.9e-158;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1550 TCATCTCTTGGTACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTAC 1609
Db 1 TCATCTCTTGGTACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTAC 60

QY 1610 TCTTTTGTGTGTGTGTGAAACAGGATCTTGTCTGTACCCAGGCTTGAGTGCAGTGG 1669
Db 61 TCTTTTGTGTGTGTGTGAAACAGGATCTTGTCTGTACCCAGGCTTGAGTGCAGTGG 120

QY 1670 TGGGATCAGACCCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCAT 1729
Db 121 TGGGATCAGACCCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCAT 180

QY 1730 CTCCTCAAGTGTCTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTT 1789
Db 181 CTCCTCAAGTGTCTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTT 240

QY 1790 GCTGTTATTTTCATGGACTAAAGGTCTGGTCACTAGCTCAGCTGGCTGCACAGCT 1849
Db 241 GCTGTTATTTTCATGGACTAAAGGTCTGGTCACTAGCTCAGCTGGCTGCACAGCT 300

QY 1850 CTAGGGGCTGTCTTAACTCAGAGTGGTGTGTGAGGCTCTGTGGCCAGAGCAGA 1909
Db 301 YTAGGGGCTGTCTTAACTCAGAGTGGTGTGTGAGGCTCTGTGGCCAGAGCAGA 360

QY 1910 CTTGCATATCTGAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACAC 1969
Db 361 CTTGCATATCTGAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACAC 420

QY 1970 TGAAGCCAACTTGTGTGGCAACCCCTCCCAACCTTCTTGCCTGGGTAGGAGGCT 2029
Db 421 TGAAGCCAACTTGTGTGGCAACCCCTCCCAACCTTCTTGCCTGGGTAGGAGGCT 480

```
QY 2030 AAAGATCACCCCTAAATTTACTATCTCTCTAGTGTGCTCTCAATTTGGCCCTCAGCAGCT 2089
Db 481 AAAGATCACCCCTAAATTTACTATCTCTCTAGTGTGCTCTCAATTTGGCCCTCAGCAGCT 540
QY 2090 CCCCAGCACCAATTCACAGGTACCCCTCTCTCTTGGCACTGCCCAAACTTGCTGTCA 2149
Db 541 CCCCAGCACCAATTCACAGGTACCCCTCTCTCTTGGCACTGCCCAAACTTGCTGTCA 600
QY 2150 A 2150
Db 601 A 601

RESULT 12
US-09-949-016-29170
; Sequence 29170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29170
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-29170

Query Match 17.0%; Score 418.6; DB 3; Length 601;
Best Local Similarity 99.8%; Pred. No. 4.2e-107;
Matches 418; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 246 GGTCCAGAAATGGCTACAAACAGTCCCTGTGATGGTTGCCGCAATGCCCTGAAATCCAAGTACT 305
Db 183 GGTCCAGAAATGGCTACAAACAGTCCCTGTGATGGTTGCCGCAATGCCCTGAAATCCAAGTACT 242
QY 306 ACAGACTTTTGTATAGGCTGAAGCTTGGGGCATCTCTAGAAAACGGTGGCCACAGCCG 365
Db 243 ACAGACTTTTGTATAGGCTGAAGCTTGGGGCATCTCTAGAAAACGGTGGCCACAGCYG 302
QY 366 GGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGTGCAGG 425
Db 303 GGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGTGCAGG 362
QY 426 ACTCCAAACAGGGGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCTGGGTGTGTGGGCA 485
Db 363 ACTCCAAACAGGGGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCTGGGTGTGTGGGCA 422
QY 486 TCTTTGGCCTACCTTCCTTCCTTCATCATCGGACTGACGGGAGCAGAGGCCACACGCT 545
Db 423 TCTTTGGCCTACCTTCCTTCCTTCATCATCGGACTGACGGGAGCAGAGGCCACACGCT 482
QY 546 TCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTCCCTGCTGGCTCATGCTGCA 605
Db 483 TCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTCCCTGCTGGCTCATGCTGCA 542
QY 606 GTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTCCCTGTGGTGTCTGGGTCTG 664
Db 543 GTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTCCCTGTGGTGTCTGGGTCTG 601

RESULT 13
US-09-949-016-126954
; Sequence 126954, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126954
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-126954

Query Match 17.0%; Score 418.6; DB 3; Length 601;
Best Local Similarity 99.8%; Pred. No. 4.2e-107;
Matches 418; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 246 GGTCCAGAAATGGCTACAAACAGTCCCTGTGATGGTTGCCGCAATGCCCTGAAATCCAAGTACT 305
Db 183 GGTCCAGAAATGGCTACAAACAGTCCCTGTGATGGTTGCCGCAATGCCCTGAAATCCAAGTACT 242
QY 306 ACAGACTTTTGTATAGGCTGAAGCTTGGGGCATCTCTAGAAAACGGTGGCCACAGCCG 365
Db 243 ACAGACTTTTGTATAGGCTGAAGCTTGGGGCATCTCTAGAAAACGGTGGCCACAGCTG 302
QY 366 GGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGTGCAGG 425
Db 303 GGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGTGCAGG 362
QY 426 ACTCCAAACAGGGGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCTGGGTGTGTGGGCA 485
Db 363 ACTCCAAACAGGGGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCTGGGTGTGTGGGCA 422
QY 486 TCTTTGGCCTACCTTCCTTCCTTCATCATCGGACTGACGGGAGCAGAGGCCACACGCT 545
Db 423 TCTTTGGCCTACCTTCCTTCCTTCATCATCGGACTGACGGGAGCAGAGGCCACACGCT 482
QY 546 TCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGGCTCATGCTGTCA 605
Db 483 TCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGGCTCATGCTGTCA 542
QY 606 GTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTCCCTGTGGTGTCTGGGTCTG 664
Db 543 GTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTCCCTGTGGTGTCTGGGTCTG 601

RESULT 13
US-09-949-016-126954
```

```
; Sequence 126954, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126954
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-126954

Query Match 17.0%; Score 418.6; DB 3; Length 601;
Best Local Similarity 99.8%; Pred. No. 4.2e-107;
Matches 418; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 246 GGTCCAGAAATGGCTACAAACAGTCCCTGTGATGGTTGCCGCAATGCCCTGAAATCCAAGTACT 305
Db 183 GGTCCAGAAATGGCTACAAACAGTCCCTGTGATGGTTGCCGCAATGCCCTGAAATCCAAGTACT 242
QY 306 ACAGACTTTTGTATAGGCTGAAGCTTGGGGCATCTCTAGAAAACGGTGGCCACAGCCG 365
Db 243 ACAGACTTTTGTATAGGCTGAAGCTTGGGGCATCTCTAGAAAACGGTGGCCACAGCYG 302
QY 366 GGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGTGCAGG 425
Db 303 GGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGTGCAGG 362
QY 426 ACTCCAAACAGGGGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCTGGGTGTGTGGGCA 485
Db 363 ACTCCAAACAGGGGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCTGGGTGTGTGGGCA 422
QY 486 TCTTTGGCCTACCTTCCTTCCTTCATCATCGGACTGACGGGAGCAGAGGCCACACGCT 545
Db 423 TCTTTGGCCTACCTTCCTTCCTTCATCATCGGACTGACGGGAGCAGAGGCCACACGCT 482
QY 546 TCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTCCCTGCTGGCTCATGCTGCA 605
Db 483 TCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTCCCTGCTGGCTCATGCTGCA 542
QY 606 GTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTCCCTGTGGTGTCTGGGTCTG 664
Db 543 GTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTCCCTGTGGTGTCTGGGTCTG 601

RESULT 14
US-09-188-930-4
; Sequence 4, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
```

; SEQ ID NO 4
; LENGTH: 311
; TYPE: DNA
; ORGANISM: Human
; US-09-188-930-4

Query Match 11.9%; Score 292.2; DB 3; Length 311;
Best Local Similarity 99.0%; Pred. No. 8.1e-72;
Matches 294; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

195 TGTGGGCATCTTTGGCCTCACCTTCGCTTCATCGGACTGGACGGGAGCACAGGC 254
537 CCACACGCTTCTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTCTGCTGG 593
255 CCACACGCTTCTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTCTGCTGG 311

Search completed: June 10, 2006, 15:26:30
Job time : 433 secs

297 CCAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGGCATCGTCCTAGAAAACGGTGG 356
15 CGAGGTACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGGCATCGTCCTAGAAAACGGTGG 74
357 CCACAGCGGGGTGTGACCTCGGTGGCTTCACTGCTCACTCTCCGATCCTCGTCTGCA 416
75 CCACAGCTGGGGTGTGACCTCGGTGGCTTCACTGCTCACTCTCCGATCCTCGTCTGCA 134
417 AGGTGAGGACTCCAAACAGCGGAAAAATGCTGCTACTCAGTTTCTCTCTCTCTGGTG 476
135 AGGTGAGGACTCCAAACAGCGGAAAAATGCTGCTACTCAGTTTCTCTCTCTGGTG 194
477 TGTGGGCATCTTTGGCCTCACCTTCGCTTCATATCGGACTGGACGGAGCACAGGC 536
195 TGTGGGCATCTTTGGCCTCACCTTCGCTTCATATCGGACTGGACGGAGCACAGGC 254
537 CCACACGCTTCTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGG 593
255 CCACACGCTTCTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGG 311

RESULT 15
US-09-312-283C-4
; Sequence 4, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 311
; TYPE: DNA
; ORGANISM: Mouse
; US-09-312-283C-4

Query Match 11.9%; Score 292.2; DB 3; Length 311;
Best Local Similarity 99.0%; Pred. No. 8.1e-72;
Matches 294; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

297 CCAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGGCATCGTCCTAGAAAACGGTGG 356
15 CGAGGTACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGGCATCGTCCTAGAAAACGGTGG 74
357 CCACAGCGGGGTGTGACCTCGGTGGCTTCACTGCTCACTCTCCGATCCTCGTCTGCA 416
75 CCACAGCTGGGGTGTGACCTCGGTGGCTTCACTGCTCACTCTCCGATCCTCGTCTGCA 134
417 AGGTGAGGACTCCAAACAGCGGAAAAATGCTGCTACTCAGTTTCTCTCTCTGGTG 476
135 AGGTGAGGACTCCAAACAGCGGAAAAATGCTGCTACTCAGTTTCTCTCTCTGGTG 194
477 TGTGGGCATCTTTGGCCTCACCTTCGCTTCATATCGGACTGGACGGGAGCACAGGC 536

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:08:58 ; Search time 3003 Seconds
(without alignments)
10049.428 Million cell updates/sec

Title: US-10-600-816-2

Perfect score: 2456

Sequence: 1 ataacagcatgaagtcgcgt.....ggcagcaaaaaaaaaa 2456

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2456	100.0	2456	6	US-10-225-567A-453
2	2456	100.0	2456	7	US-10-269-909-63
3	2456	100.0	2456	7	US-10-269-909-64
4	2456	100.0	2456	7	US-10-295-027-619
5	2456	100.0	2456	8	US-10-600-816-2
6	2456	100.0	2456	9	US-10-775-920-10
7	2456	100.0	2456	10	US-10-936-626-40
8	2456	100.0	2456	10	US-10-938-061-40
9	2456	100.0	2456	16	US-11-169-041-32
10	2448	99.7	2456	8	US-10-600-816-18
11	2443	99.5	4239	6	US-10-198-846-10424
12	2439.8	99.3	2446	9	US-10-775-920-11
13	2302	93.7	2302	7	US-10-224-289-3
14	2302	93.7	2302	8	US-10-240-425-405
15	2302	93.7	2302	9	US-10-775-920-9
16	2302	93.7	2302	10	US-10-510-507-2
17	2286	93.1	2316	6	US-10-176-847-59

18	2286	93.1	2316	15	US-11-080-991-59	Sequence 59, Appl
19	2274.8	92.6	2297	9	US-10-775-920-13	Sequence 13, Appl
20	2268.2	92.4	2305	9	US-10-775-920-13	Sequence 12, Appl
21	2260.4	92.0	2593	7	US-10-264-049-834	Sequence 834, App
22	1601.8	65.2	1619	7	US-10-224-289-5	Sequence 5, Appl1
23	1601.8	65.2	1619	9	US-10-935-190-43	Sequence 43, Appl1
24	1441.4	58.7	1460	10	US-10-936-626-64	Sequence 64, Appl
25	1441.4	58.7	1460	10	US-10-938-061-64	Sequence 64, Appl
26	1400	57.0	1400	13	US-11-060-756-2418	Sequence 2418, Ap
27	1400	57.0	1400	13	US-11-060-756-6690	Sequence 6690, Ap
28	1126	45.8	1212	3	US-09-866-050A-249	Sequence 249, App
29	1126	45.8	1212	6	US-10-152-661-210	Sequence 210, App
30	1082.6	44.1	1114	3	US-09-978-360A-210	Sequence 210, App
31	1082.6	44.1	1114	5	US-09-978-360A-210	Sequence 210, App
32	1071	43.6	1071	9	US-10-712-615-134	Sequence 134, App
33	1067.8	43.5	1788	10	US-10-505-486-196	Sequence 196, App
34	1031.6	42.0	1228	7	US-10-313-542-223	Sequence 223, App
C 35	932.6	38.0	1000	3	US-09-864-761-19238	Sequence 19238, A
C 36	787.6	32.1	948	3	US-09-864-761-2510	Sequence 2510, Ap
C 37	702.4	28.6	774	6	US-10-106-698-2079	Sequence 2079, Ap
C 38	661.2	26.9	1934	7	US-10-224-289-1	Sequence 1, Appli
39	602	24.5	715	3	US-09-969-034-1270	Sequence 1270, Ap
C 40	540.4	22.0	620	3	US-09-969-034-2222	Sequence 2222, Ap
C 41	538.4	21.9	552	6	US-10-066-543-1811	Sequence 1811, Ap
C 42	497	20.2	497	6	US-10-066-543-1937	Sequence 1937, Ap
C 43	496.6	20.2	562	3	US-09-969-034-1808	Sequence 1808, Ap
C 44	477.2	19.4	519	6	US-10-052-283-396	Sequence 396, App
C 45	467	19.0	468	3	US-09-998-598-2451	Sequence 2451, Ap

ALIGNMENTS

RESULT 1

US-10-225-567A-453
; Sequence 453, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 453
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-453

Query Match	100.0%	Score 2456;	DB 6;	Length 2456;
Best Local Similarity	100.0%;	Pred. No. 0.		
Matches 2456;	Conservative 0;	Mismatches 0;	Gaps 0;	
QY	1	ATAACAGCATGAAGTCGCGTGGAACTGGAACTAGGCGGTGCTCTCCCTCGACCCCTCCCC	60	
Db	1	ATAACAGCATGAAGTCGCGTGGAACTGGAACTAGGCGGTGCTCTCCCTCGACCCCTCCCC	60	
QY	61	TCCTTGTCCTCTGCTCACCCCTCGCTTCCTCCCTCCGCGAGGGCGGCTTTATA	120	
Db	61	TCCTTGTCCTCTGCTCACCCCTCGCTTCCTCCCTCCGCGAGGGCGGCTTTATA	120	
QY	121	ACAACTGCTCAGATGCGAGGGGGGATAGCTGTCCAAAGTCTCCCGGAGCTAGGAG	180	
Db	121	ACAACTGCTCAGATGCGAGGGGGGATAGCTGTCCAAAGTCTCCCGGAGCTAGGAG	180	
QY	181	CTCGCCTGTCCTCTTTCGCGCGGGAAGACAGCACCAGTTTCACGGCCAAACGCTTGGC	240	

Db	181	CTCGGCTGTGCCCCTCTTGGCGCGGGAAGAGAGACCAAGTTACGGCCAAACGCGCTTGGC	240
Qy	241	ACTAGGGTCCAGAAATGGCTACAAAGTCCCTGATGGTTGCCGCAATGGGCTGAATCCAA	300
Db	241	ACTAGGGTCCAGAAATGGCTACAAAGTCCCTGATGGTTGCCGCAATGGGCTGAATCCAA	300
Qy	301	GTACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGCAATCGTCTTAGAAACCGGTGGCCAC	360
Db	301	GTACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGCAATCGTCTTAGAAACCGGTGGCCAC	360
Qy	361	AGCCGGGGTGTGACCTCGGTGGCCCTTATGCTCTCACTCTCCGATCCCTGTCGAAGT	420
Db	361	AGCCGGGGTGTGACCTCGGTGGCCCTTATGCTCTCACTCTCCGATCCCTGTCGAAGT	420
Qy	421	GCAGAGACTCCAAAGGCGGAAAAATGCTGCTACTCAGTTTCTCTTCTCTCTCGGTGTGTT	480
Db	421	GCAGAGACTCCAAAGGCGGAAAAATGCTGCTACTCAGTTTCTCTTCTCTCTCGGTGTGTT	480
Qy	481	GGGCATCTTTGGCCTCACTTCCGCTTCATATCGGAAGTGAAGCGGAGACAGGGGCCAC	540
Db	481	GGGCATCTTTGGCCTCACTTCCGCTTCATATCGGAAGTGAAGCGGAGACAGGGGCCAC	540
Qy	541	AGCCTTCTCTCTTTGGGATCCTCTTTTTCATCTGCTTCTCTCTCTGCTGCTGCTCATGC	600
Db	541	AGCCTTCTCTCTTTGGGATCCTCTTTTTCATCTGCTTCTCTCTCTGCTGCTGCTCATGC	600
Qy	601	TGTCAGTCTGACAAAGCTCGTCCGGGGAGGAGGCCCTTTCCCTGTTGGTGAATCTGGG	660
Db	601	TGTCAGTCTGACAAAGCTCGTCCGGGGAGGAGGCCCTTTCCCTGTTGGTGAATCTGGG	660
Qy	661	TCTGGCGGTGGGCTTCAAGCTAGTCCAGGATGTTATCGCTATTTGAATATATTTGCTCGAC	720
Db	661	TCTGGCGGTGGGCTTCAAGCTAGTCCAGGATGTTATCGCTATTTGAATATATTTGCTCGAC	720
Qy	721	CATGATAGGACCAAGTCAATGCTTTTCTCAGCTTTCCGCTCTCGTCGCAATGAAGA	780
Db	721	CATGATAGGACCAAGTCAATGCTTTTCTCAGCTTTCCGCTCTCGTCGCAATGAAGA	780
Qy	781	CTTTGTCTCTCTGCTCACCCTAGCTCTCTCTTTGATGGGCTGACCTTCTCATGTCCTC	840
Db	781	CTTTGTCTCTCTGCTCACCCTAGCTCTCTCTTTGATGGGCTGACCTTCTCATGTCCTC	840
Qy	841	CTTCACTTCTGTGTTCTTCACTGCTTCAAGGCTGGAAGACATGGGGCCCAATCTACTCAC	900
Db	841	CTTCACTTCTGTGTTCTTCACTGCTTCAAGGCTGGAAGACATGGGGCCCAATCTACTCAC	900
Qy	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACTGCTGCTCATGCTTCTCGACTT	960
Db	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACTGCTGCTCATGCTTCTCGACTT	960
Qy	961	TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTTGGCTGGCCAAATGGCTGGGTGTT	1020
Db	961	TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTTGGCTGGCCAAATGGCTGGGTGTT	1020
Qy	1021	CCTGTTGGCTTATGTTAGTCCGAGTTTGGCTGTCTCAAAAGCAACGAAACCCCATGGA	1080
Db	1021	CCTGTTGGCTTATGTTAGTCCGAGTTTGGCTGTCTCAAAAGCAACGAAACCCCATGGA	1080
Qy	1081	TTATCTGTTGAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGGA	1140
Db	1081	TTATCTGTTGAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGGA	1140
Qy	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTTGAAGACAGAGGACACGCT	1200
Db	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTTGAAGACAGAGGACACGCT	1200
Qy	1201	CTATGCCCCCTATTTCCACATTTTCAGCTCGAGAACCGCTCCCCAAAAGGAAATTTCTC	1260
Db	1201	CTATGCCCCCTATTTCCACATTTTCAGCTCGAGAACCGCTCCCCAAAAGGAAATTTCTC	1260
Qy	1261	CATCCACGGGCCAGGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG	1320

Db	1261	CATCCACGGGCCACGCTTGGCCGAGGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG	1320
Qy	1321	CAGCTAACTCTGTCTCTGAAGAGTGGGACAAAATGCAGCCGGGCGGAGACTTCTAGCGGAGC	1380
Db	1321	CAGCTAACTCTGTCTCTGAAGAGTGGGACAAAATGCAGCCGGGCGGAGACTTCTAGCGGAGC	1380
Qy	1381	TCAAAAGGATGTGGGCGAAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACTACGGGA	1440
Db	1381	TCAAAAGGATGTGGGCGAAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACTACGGGA	1440
Qy	1441	ACAGTTTGCCTCCCTCCAGCCCTCAACACAAATCTTCCATCTCGGGCTGATGTGGGT	1500
Db	1441	ACAGTTTGCCTCCCTCCAGCCCTCAACACAAATCTTCCATCTCGGGCTGATGTGGGT	1500
Qy	1501	AGTAAGACTCCTAGTTCTTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTCTGG	1560
Db	1501	AGTAAGACTCCTAGTTCTTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTCTGG	1560
Qy	1561	ATACTTCTTTTAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGTGTT	1620
Db	1561	ATACTTCTTTTAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGTGTT	1620
Qy	1621	GTTTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGGTGCATCACAG	1680
Db	1621	GTTTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGGTGCATCACAG	1680
Qy	1681	CCCAGTGCAGCTCCGACCACTGTCTCAAGCAATCTCCCATCTCCATACTCTCCAAAGTG	1740
Db	1681	CCCAGTGCAGCTCCGACCACTGTCTCAAGCAATCTCCCATCTCCATACTCTCCAAAGTG	1740
Qy	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGGCTTAGGCCCTTAATCTTGTCTGTTATTTT	1800
Db	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGGCTTAGGCCCTTAATCTTGTCTGTTATTTT	1800
Qy	1801	CCATCGACTAAAGTCTGGTCTATCTGAGCTACGCTGCTGCATCACAGCTCTTAGGGCCCTG	1860
Db	1801	CCATCGACTAAAGTCTGGTCTATCTGAGCTACGCTGCTGCATCACAGCTCTTAGGGCCCTG	1860
Qy	1861	CTTCTCTAACTACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCTATATCT	1920
Db	1861	CTTCTCTAACTACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCTATATCT	1920
Qy	1921	GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGTAATCTACACTGGAAGCCAAAC	1980
Db	1921	GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGTAATCTACACTGGAAGCCAAAC	1980
Qy	1981	TTGCTGGCACCCCGCTCCCCAAACCTTCTTCCCTGGGTAGGAGAGGCTTAAAGATCACCC	2040
Db	1981	TTGCTGGCACCCCGCTCCCCAAACCTTCTTCCCTGGGTAGGAGAGGCTTAAAGATCACCC	2040
Qy	2041	TAAATTTACTCATCTCTTAGTGTGCCTTCAATTTGGGCTCAGCAGCTTCCCGAGCAACA	2100
Db	2041	TAAATTTACTCATCTCTTAGTGTGCCTTCAATTTGGGCTCAGCAGCTTCCCGAGCAACA	2100
Qy	2101	ATTTCAGGTCACCCCTCTCTTCTTTCGACTGTGCCCAACTTGTCTCAATTTCCGAGATC	2160
Db	2101	ATTTCAGGTCACCCCTCTCTTCTTTCGACTGTGCCCAACTTGTCTCAATTTCCGAGATC	2160
Qy	2161	TAAATCTCCCTTACGCTCTGCGAGGAAATTTCTTTTCAGACCTCTACAGCAAGCCGGTTG	2220
Db	2161	TAAATCTCCCTTACGCTCTGCGAGGAAATTTCTTTTCAGACCTCTACAGCAAGCCGGTTG	2220
Qy	2221	CTTCTTGTGAGGAAATTTGTAGATCAATTTCAATTTCTGAGGCTGATCTTCT	2280
Db	2221	CTTCTTGTGAGGAAATTTGTAGATCAATTTCAATTTCTGAGGCTGATCTTCT	2280
Qy	2281	CTCATCTTGCACCCCAACCTCTGTAATAGATTTTACCAGATTTAGGCTGCAATCTGTA	2340
Db	2281	CTCATCTTGCACCCCAACCTCTGTAATAGATTTTACCAGATTTAGGCTGCAATCTGTA	2340
Qy	2341	GTGGGCAATGGTCTCTCTTAAATGGAGGAGTGTTCATTTGATAATTAAGTTATTTCACCTGAGTAT	2400
Db	2341	GTGGGCAATGGTCTCTCTTAAATGGAGGAGTGTTCATTTGATAATTAAGTTATTTCACCTGAGTAT	2400


```
QY 1681 CCCAGTGCAGCCCTCGACCAACCTGTGTCTAAGCAATCCTCCCATCTCCCAAGTG 1740
Db 1681 CCCAGTGCAGCCCTCGACCAACCTGTGTCTAAGCAATCCTCCCATCTCCCAAGTG 1740
QY 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGCTAGGCCCTTAATCTTGTGTTATTTT 1800
Db 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGCTAGGCCCTTAATCTTGTGTTATTTT 1800
QY 1801 CCATGGAATAAAGGTCTGTCTCATCTGAGCTCAGCTGGCTCACACAGCTCTTAGGGCCCTG 1860
Db 1801 CCATGGAATAAAGGTCTGTCTCATCTGAGCTCAGCTGGCTCACACAGCTCTTAGGGCCCTG 1860
QY 1861 CTCCTCTAACTACAGTGGGTTTGTAGGCTCTGTGGCCACAGAGACACCTGCATATCT 1920
Db 1861 CTCCTCTAACTACAGTGGGTTTGTAGGCTCTGTGGCCACAGAGACACCTGCATATCT 1920
QY 1921 GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCCTGAATCTACACTGGAAGCCAC 1980
Db 1921 GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCCTGAATCTACACTGGAAGCCAC 1980
QY 1981 TTGCTGGCACCCCGCTCCGCCAACCTTCTTGCCTGGGTAGGAGGCTTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCCGCTCCGCCAACCTTCTTGCCTGGGTAGGAGGCTTAAAGATCACCC 2040
QY 2041 TAAATTTACTCATCTCTTAGTGTGCTCTCAATTTGGGCTCAGCAGCTCCCGACCA 2100
Db 2041 TAAATTTACTCATCTCTTAGTGTGCTCTCAATTTGGGCTCAGCAGCTCCCGACCA 2100
QY 2101 ATTACAGTCAACCTCTCTTCTTGCAGTGTCCCAAACTTGTCTCAATTTCCGAGATC 2160
Db 2101 ATTACAGTCAACCTCTCTTCTTGCAGTGTCCCAAACTTGTCTCAATTTCCGAGATC 2160
QY 2161 TAATCTCCCTTACGCTCTGCCAGGAATTTCTTTTCAGACCTCTACAGCAAGCCCGGTTG 2220
Db 2161 TAATCTCCCTTACGCTCTGCCAGGAATTTCTTTTCAGACCTCTACAGCAAGCCCGGTTG 2220
QY 2221 CTCCTTGTGAGAGAAATTTGTAGATCTCTCACTTCAAATTTCTGGGCTGATCTCT 2280
Db 2221 CTCCTTGTGAGAGAAATTTGTAGATCTCTCACTTCAAATTTCTGGGCTGATCTCT 2280
QY 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTTACCGATTTACGGCTGCAATCTGTAA 2340
Db 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTTACCGATTTACGGCTGCAATCTGTAA 2340
QY 2341 GTGGCATGGTCTCCTAATGAGGAGTGTTCATTTGATTAATTAAGTTATTCACTCAGTAT 2400
Db 2341 GTGGCATGGTCTCCTAATGAGGAGTGTTCATTTGATTAATTAAGTTATTCACTCAGTAT 2400
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456
```

```
RESULT 3
US-10-269-909-64
; Sequence 64, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARBANI, PEDRAM
; APPLICANT: IACOBUIZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 64
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-64
```

```
Query Match 100.0%; Score 2456; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATAAACAGATGAAGTGCCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC 60
Db 1 ATAAACAGATGAAGTGCCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC 60
QY 61 TCTTTGTCCCTCTGCTACCCCTCGCTCCCTCCCTCCCTCCCGGAGGCGCGCTTTATA 120
Db 61 TCTTTGTCCCTCTGCTACCCCTCGCTCCCTCCCTCCCTCCCGGAGGCGCGCTTTATA 120
QY 121 ACAACTGCTCAGAGTGCAGGGGGGATAGCTGTCCNAGGTCTCCCGCAGCACTGAGGAG 180
Db 121 ACAACTGCTCAGAGTGCAGGGGGGATAGCTGTCCNAGGTCTCCCGCAGCACTGAGGAG 180
QY 181 CTCGCTGTGCTCCCTCTTTCGCGCGGGAAGCAGCACCAGTTTCAGGCCAAAGCCTTGGC 240
Db 181 CTCGCTGTGCTCCCTCTTTCGCGCGGGAAGCAGCACCAGTTTCAGGCCAAAGCCTTGGC 240
QY 241 ACTAGGGTCCAGAATGGCTTACAACTGCTGATGGTTGGCCGCAATGGCTGAAATCCAA 300
Db 241 ACTAGGGTCCAGAATGGCTTACAACTGCTGATGGTTGGCCGCAATGGCTGAAATCCAA 300
QY 301 GTACTACAGACTTTGTGATAAGGCTTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC 360
Db 301 GTACTACAGACTTTGTGATAAGGCTTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC 360
QY 361 AGCCGGGTTGTGACCTCGGTGGCTTCAATGCTCACTCTCCGATCTCTCGTCTGCAAGGT 420
Db 361 AGCCGGGTTGTGACCTCGGTGGCTTCAATGCTCACTCTCCGATCTCTCGTCTGCAAGGT 420
QY 421 GCAGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTTCTCTCTGGGTGTGT 480
Db 421 GCAGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTTCTCTCTGGGTGTGT 480
QY 481 GGGCATCTTGGCCCTCACTTCGCTTCACTCATCGSACTGAGCGGAGCAGAGGGCCAC 540
Db 481 GGGCATCTTGGCCCTCACTTCGCTTCACTCATCGSACTGAGCGGAGCAGAGGGCCAC 540
QY 541 AGCTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Db 541 AGCTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
QY 601 TGTGAGTGCACAAAGCTGCTCCGGGGAGGAAGCCCTTTCTGTTGGTGAATCTGGG 660
Db 601 TGTGAGTGCACAAAGCTGCTCCGGGGAGGAAGCCCTTTCTGTTGGTGAATCTGGG 660
QY 661 TCTGGCCGTGGGCTTTCAGCTAGTCCAGATGTATCGCTATTGATATATTGCTCTGAC 720
Db 661 TCTGGCCGTGGGCTTTCAGCTAGTCCAGATGTATCGCTATTGATATATTGCTCTGAC 720
QY 721 CATGAATAGGACCAACAGTCAATGTCTTTTCTGAGCTTTCCGCTCTCGTCCGAATGAAGA 780
Db 721 CATGAATAGGACCAACAGTCAATGTCTTTTCTGAGCTTTCCGCTCTCGTCCGAATGAAGA 780
QY 781 CTTTGTCTCTGCTCACTGCTCTCTTCTTGTGATGGCGTGACCTTCTCATGTCTCTC 840
Db 781 CTTTGTCTCTGCTCACTGCTCTCTTCTTGTGATGGCGTGACCTTCTCATGTCTCTC 840
QY 841 CTTTCACTTCTGTGTTTCTTCCAGGGCTGGAAGACATGGGGCCCATCTACCTCAC 900
Db 841 CTTTCACTTCTGTGTTTCTTCCAGGGCTGGAAGACATGGGGCCCATCTACCTCAC 900
QY 901 GATGCTCTCTCCATTTGCCATCTGGGTGGCCTGGATCACCTCTCATGCTCTCTGACTT 960
Db 901 GATGCTCTCTCCATTTGCCATCTGGGTGGCCTGGATCACCTCTCATGCTCTCTGACTT 960
```

QY 961 TGACCCGAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT 1020
DB 961 TGACCCGAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT 1020
QY 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
DB 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
QY 1081 TTATCTGTTGAGGATGCTTCTGTAACCTCAACTCGTGAAGAGCTATGTTGTTGGA 1140
DB 1081 TTATCTGTTGAGGATGCTTCTGTAACCTCAACTCGTGAAGAGCTATGTTGTTGGA 1140
QY 1141 GAACAGAGCTACTCTCAAGAGGAATCACTCAAGGTTTTTGAAGAGACAGGGACACGCT 1200
DB 1141 GAACAGAGCTACTCTCAAGAGGAATCACTCAAGGTTTTTGAAGAGACAGGGACACGCT 1200
QY 1201 CTATGCCCTTATTCACACATTTTCAGCTGCAAGACCAAGCTCCCAAGAAATTTCTC 1260
DB 1201 CTATGCCCTTATTCACACATTTTCAGCTGCAAGACCAAGCTCCCAAGAAATTTCTC 1260
QY 1261 CATCCACGGGGCCAGCTTGGCCGAGCCCTTCAAGAGACTATGAAGTAAAGAAAGGG 1320
DB 1261 CATCCACGGGGCCAGCTTGGCCGAGCCCTTCAAGAGACTATGAAGTAAAGAAAGGG 1320
QY 1321 CAGCTAACTCTGCTGAAGAGTGGACAAATCGAGCGGGCGCAGATCTAGCGGGAGC 1380
DB 1321 CAGCTAACTCTGCTGAAGAGTGGACAAATCGAGCGGGCGCAGATCTAGCGGGAGC 1380
QY 1381 TCAAGGGATGTGGGGGAAATCTTGAGTCTTCTGAGAAAATGTACAAGACACTACGGGA 1440
DB 1381 TCAAGGGATGTGGGGGAAATCTTGAGTCTTCTGAGAAAATGTACAAGACACTACGGGA 1440
QY 1441 ACAGTTTGCTCCTCCAGCTCAACACAAATCTTCCATGCTGGGGCTGATGGGCT 1500
DB 1441 ACAGTTTGCTCCTCCAGCTCAACACAAATCTTCCATGCTGGGGCTGATGGGCT 1500
QY 1501 AGTAAGACTCCAGTCTTAGAGGCGCTGATGATTTTTTTTTTTTGTCTCATCTTTGG 1560
DB 1501 AGTAAGACTCCAGTCTTAGAGGCGCTGATGATTTTTTTTTTTTGTCTCATCTTTGG 1560
QY 1561 ATACTTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTTTTGT 1620
DB 1561 ATACTTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTTTTGT 1620
QY 1621 GTTTTTTGAACAGGATCTTGCTGTACCCAGGCTTGAGTCAGTGGTGCATCACAG 1680
DB 1621 GTTTTTTGAACAGGATCTTGCTGTACCCAGGCTTGAGTCAGTGGTGCATCACAG 1680
QY 1681 CCCAGTGCAGCTCGACCCCTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAAGTG 1740
DB 1681 CCCAGTGCAGCTCGACCCCTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAAGTG 1740
QY 1741 CTGGGATGACAGGCGTGAGCCACAGCTCCCAGCTAGGCGCTTAAATCTTGCTGTTATTT 1800
DB 1741 CTGGGATGACAGGCGTGAGCCACAGCTCCCAGCTAGGCGCTTAAATCTTGCTGTTATTT 1800
QY 1801 CCATGACTAAAGGTCTGTCATCTGAGCTCAAGCTGGCTCAACAGCTTAGGGGCGCTG 1860
DB 1801 CCATGACTAAAGGTCTGTCATCTGAGCTCAAGCTGGCTCAACAGCTTAGGGGCGCTG 1860
QY 1861 CTCCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCACAGCAGACCTGCATATCT 1920
DB 1861 CTCCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCACAGCAGACCTGCATATCT 1920
QY 1921 GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCAAC 1980
DB 1921 GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCAAC 1980
QY 1981 TTGCTGACCCCCGCTCCCAACCTTCTTGCTGGTAGGAGGCTAAAGATCACCC 2040
DB 1981 TTGCTGACCCCCGCTCCCAACCTTCTTGCTGGTAGGAGGCTAAAGATCACCC 2040

QY 2041 TAAATTTACTCATCTCTAGTGTGCTCCTCAATTTGGGCTCAGAGCTCCCCAGCACCA 2100
DB 2041 TAAATTTACTCATCTCTAGTGTGCTCCTCAATTTGGGCTCAGAGCTCCCCAGCACCA 2100
QY 2101 ATTACAGGTCAACCCCTCTCTTTGCACTGTCTCCCAACTTGTCTGCAATTCGAGATC 2160
DB 2101 ATTACAGGTCAACCCCTCTCTTTGCACTGTCTCCCAACTTGTCTGCAATTCGAGATC 2160
QY 2161 TAAATCTCCCTTACGCTCTGCCAGGAATTTCTTTTCAGACCTCACTAGCACAAAGCCGTTG 2220
DB 2161 TAAATCTCCCTTACGCTCTGCCAGGAATTTCTTTTCAGACCTCACTAGCACAAAGCCGTTG 2220
QY 2221 CTCCTTTGTGAGAGAAATTTGTAGATCATTTCTCACTTTCAAAATTCCTGGGCTGATCTCT 2280
DB 2221 CTCCTTTGTGAGAGAAATTTGTAGATCATTTCTCACTTTCAAAATTCCTGGGCTGATCTCT 2280
QY 2281 CTCATCTTGCAACCCCACTCTGTAAATAGATTTACGCAATTTACGGCTGCATCTGTAA 2340
DB 2281 CTCATCTTGCAACCCCACTCTGTAAATAGATTTACGCAATTTACGGCTGCATCTGTAA 2340
QY 2341 GTGGGATGCTCTCTTAATGGAGGAGTTCATTGTATATAAGTTATTACCTGAGTAT 2400
DB 2341 GTGGGATGCTCTCTTAATGGAGGAGTTCATTGTATATAAGTTATTACCTGAGTAT 2400
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456
DB 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456

RESULT 4

US-10-295-027-619
; Sequence 619, Application US/10295027
; Publication No. US200302350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 619

; LENGTH: 2456			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-295-027-619			
Query Match 100.0%; Score 2456; DB 7; Length 2456;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATAACAGCATGAAGTCCGCTGGAACCTGGAATAGGCGGTGTCCTCTCCCTCGACCCCTCCCC	60
DB	1	ATPAAAGCATGAAGTCCGCTGGAACCTGGAATAGGCGGTGTCCTCTCCCTCGACCCCTCCCC	60
QY	61	TCCTGTCCCTCTGCTCAACCCCTCGCTTCCCTCCCTCCGCGAGGCGCCCTTTATA	120
DB	61	TCCTGTCCCTCTGCTCAACCCCTCGCTTCCCTCCCTCCGCGAGGCGCCCTTTATA	120
QY	121	ACAACTGCTCAGAGTCGAGGGGGGATAGCTGTCCAAGGTCTCCCGAGCACTGAGGAG	180
DB	121	ACAACTGCTCAGAGTCGAGGGGGGATAGCTGTCCAAGGTCTCCCGAGCACTGAGGAG	180
QY	181	CTCGCTGCTGCCCTCTTCGCGCGGGAAGCAGCACCAAGTTTCACGGCCAAAGCCCTTGGC	240
DB	181	CTCGCTGCTGCCCTCTTCGCGCGGGAAGCAGCACCAAGTTTCACGGCCAAAGCCCTTGGC	240
QY	241	ACTAGGGTCCAGATGGCTAACACAGTCCCTGATGGTTGCCCAATGGCCTGAATCCAA	300
DB	241	ACTAGGGTCCAGATGGCTAACACAGTCCCTGATGGTTGCCCAATGGCCTGAATCCAA	300
QY	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATCGTCTAGAAAACGGTGGCCAC	360
DB	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATCGTCTAGAAAACGGTGGCCAC	360
QY	361	AGCGGGGTGTGACCTCGGTGGCCCTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420
DB	361	AGCGGGGTGTGACCTCGGTGGCCCTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420
QY	421	GAAGACTCCAAAGCGGAAAATGCTGCTACTCAGATTTCTCTTCTCTCGGTGTGTT	480
DB	421	GAAGACTCCAAAGCGGAAAATGCTGCTACTCAGATTTCTCTTCTCTCGGTGTGTT	480
QY	481	GGGCATCTTTGGCTTCACTTCCCTTCACTATCGGACTGGAGCGGAGCACAGGGGCCAC	540
DB	481	GGGCATCTTTGGCTTCACTTCCCTTCACTATCGGACTGGAGCGGAGCACAGGGGCCAC	540
QY	541	AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTCTGGCTCATGC	600
DB	541	AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTCTGGCTCATGC	600
QY	601	TGTCAGTGTGACCAAGCTCGTCGGGGAGGAAGCCCTTTCCTGTTGGTGAATCTGGG	660
DB	601	TGTCAGTGTGACCAAGCTCGTCGGGGAGGAAGCCCTTTCCTGTTGGTGAATCTGGG	660
QY	661	TCGCGCGTGGGCTTCAAGCTAGTCAGAGATGTTATCGCTATTGAATATTGTCCTGAC	720
DB	661	TCGCGCGTGGGCTTCAAGCTAGTCAGAGATGTTATCGCTATTGAATATTGTCCTGAC	720
QY	721	CATGAATAGGACCAACGTCATGTCCTGAGCTTTCCGCTCTCGTCGCAATGAAGA	780
DB	721	CATGAATAGGACCAACGTCATGTCCTGAGCTTTCCGCTCTCGTCGCAATGAAGA	780
QY	781	CTTTGTCTCTGCTCACTAGTCCTCTTCTGATGGCGTGAACCTTCTCATGTCCTC	840
DB	781	CTTTGTCTCTGCTCACTAGTCCTCTTCTGATGGCGTGAACCTTCTCATGTCCTC	840
QY	841	CTTCACTTCTGTTGTTCTTCAAGGCTGGAAGACATGGGGCCACATCTACCTAC	900
DB	841	CTTCACTTCTGTTGTTCTTCAAGGCTGGAAGACATGGGGCCACATCTACCTAC	900
QY	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCCTGGATCAACCTGCTCATGCTTCTGACTT	960
DB	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCCTGGATCAACCTGCTCATGCTTCTGACTT	960
QY	961	TGACCGCAGTGGGATGACACCATCTCTAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT	1020
DB	961	TGACCGCAGTGGGATGACACCATCTCTAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT	1020
QY	1021	CTGTGTGGCTTATGTAGTCCGAGTTTGGCTGTCTCAAAAAGCAACGAAACCCCATGGA	1080
DB	1021	CTGTGTGGCTTATGTAGTCCGAGTTTGGCTGTCTCAAAAAGCAACGAAACCCCATGGA	1080
QY	1081	TTATCTGTGTAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA	1140
DB	1081	TTATCTGTGTAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA	1140
QY	1141	GAAAGAGCTTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGAGAGAGCTATGGTGTGA	1200
DB	1141	GAAAGAGCTTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGAGAGAGCTATGGTGTGA	1200
QY	1201	CTATGCCCTTATTCACACATTTTTCAGCTGAGAACCAAGCTTCCCAAAAGAAATTTCTC	1260
DB	1201	CTATGCCCTTATTCACACATTTTTCAGCTGAGAACCAAGCTTCCCAAAAGAAATTTCTC	1260
QY	1261	CATCCACAGGGCCCGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG	1320
DB	1261	CATCCACAGGGCCCGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG	1320
QY	1321	CAGCTAACTCTGCTCAAGAGTGGGAACAAATGAGCCGGGCGGAGACTAGCGGGAGC	1380
DB	1321	CAGCTAACTCTGCTCAAGAGTGGGAACAAATGAGCCGGGCGGAGACTAGCGGGAGC	1380
QY	1381	TCAAAAGGATGAGGGCGAAATCTTGAGTCTTCTGAGAAAACCTGTAACAAGACTACGGGA	1440
DB	1381	TCAAAAGGATGAGGGCGAAATCTTGAGTCTTCTGAGAAAACCTGTAACAAGACTACGGGA	1440
QY	1441	ACAGTTGCTCCCTCCAGCCTCAACCAATTTCTTCCAGCTGGGGCTGATGGGCT	1500
DB	1441	ACAGTTGCTCCCTCCAGCCTCAACCAATTTCTTCCAGCTGGGGCTGATGGGCT	1500
QY	1501	AGTAAGACTCAGTTCTTAGAGGCGCTGATGATTTTTTTTTTTTTTTTGTCTCATCTTTGG	1560
DB	1501	AGTAAGACTCAGTTCTTAGAGGCGCTGATGATTTTTTTTTTTTTTTTGTCTCATCTTTGG	1560
QY	1561	ATACCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGATTTAGACCCCTTACTCTTTTGTGTT	1620
DB	1561	ATACCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGATTTAGACCCCTTACTCTTTTGTGTT	1620
QY	1621	GTTTTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAAGTGTGCGATCACAG	1680
DB	1621	GTTTTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAAGTGTGCGATCACAG	1680
QY	1681	CCAGTGCAGCTCGACCACTGCTCAAGCAATCCTCCATCTCCATCTCCCAAGTG	1740
DB	1681	CCAGTGCAGCTCGACCACTGCTCAAGCAATCCTCCATCTCCATCTCCCAAGTG	1740
QY	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGCCTAGGCCCTTAACTCTTGTGTTATTTT	1800
DB	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGCCTAGGCCCTTAACTCTTGTGTTATTTT	1800
QY	1801	CCATGCACTAAAGGTCTGGTCACTGAGCTCACGCTGGCTCACACAGCTTAGGGGCTG	1860
DB	1801	CCATGCACTAAAGGTCTGGTCACTGAGCTCACGCTGGCTCACACAGCTTAGGGGCTG	1860
QY	1861	CTCCTCTAACTCAGAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT	1920
DB	1861	CTCCTCTAACTCAGAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT	1920
QY	1921	GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGGAATCTACACTGGAAGCCAAC	1980
DB	1921	GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGGAATCTACACTGGAAGCCAAC	1980
QY	1981	TTGCTGGACCCCGCTCCCAACCCCTTCTTGGCTGGGATGAGAGGCTTAAAGATCACCC	2040
DB	1981	TTGCTGGACCCCGCTCCCAACCCCTTCTTGGCTGGGATGAGAGGCTTAAAGATCACCC	2040
QY	2041	TAAATTTACTCTCTTAGTGCTCTCAATTTGGGCTCAGCAGCTCCCGAGCACA	2100


```
Db 2041 TAAATTTACTCATCTCTCTAGTCTGCTTGCCTCACATTGGGCTCAGCAGCTCCCCAGCACCA 2100
Qy 2101 ATTACAGGTACCCCTCTCTCTTGTGACTGTGCCAAACTTGCTGTAATCCGAGATC 2160
Db 2101 ATTACAGGTACCCCTCTCTCTTGTGACTGTGCCAAACTTGCTGTAATCCGAGATC 2160
Qy 2161 TAAATCTCCCTACAGTCTGTGCCAGGAATCTTTTCAGACCTCAGTACACAGCCCGGTG 2220
Db 2161 TAAATCTCCCTACAGTCTGTGCCAGGAATCTTTTCAGACCTCAGTACACAGCCCGGTG 2220
Qy 2221 CTCCTTGTTCAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
Db 2221 CTCCTTGTTCAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
Qy 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTACCGCATTTACGGCTGCATTCTGTAA 2340
Db 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTACCGCATTTACGGCTGCATTCTGTAA 2340
Qy 2341 GTGGGATGCTCTCCTAATGGAGAGTGTTCATTGTATATAAGTTATTCACCTGAGTAT 2400
Db 2341 GTGGGATGCTCTCCTAATGGAGAGTGTTCATTGTATATAAGTTATTCACCTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGTGGCCACTCTTTTCATGTGTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGTGGCCACTCTTTTCATGTGTGGTGGCAGCAAAAAA 2456

RESULT 5
US-10-600-816-2
; Sequence 2, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RA13, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-2

Query Match 100.0%; Score 2456; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATAAACAGCATGAAGTCCGTGGAACTGGAATAGCGGTGCTCTCCCTCGACCCCTCCCC 60
Db 1 ATAAACAGCATGAAGTCCGTGGAACTGGAATAGCGGTGCTCTCCCTCGACCCCTCCCC 60
Qy 61 TCCTTGTCCCTCTGCTCACCCCTCGTTCCTCCCTCCGCGGAGGCGCGCTTTATA 120
Db 61 TCCTTGTCCCTCTGCTCACCCCTCGTTCCTCCCTCCGCGGAGGCGCGCTTTATA 120
Qy 121 ACAATGCTCAGAGTCGAGGGCGGAGTACTGTCCAAGTCTCCCCAGCACTGAGGAG 180
Db 121 ACAATGCTCAGAGTCGAGGGCGGAGTACTGTCCAAGTCTCCCCAGCACTGAGGAG 180
Qy 181 CTCGCTGCTGCCTCTTGGCGGGGAGCAGCACAAGTTTACCGCCCAACGCTTGGC 240
Db 181 CTCGCTGCTGCCTCTTGGCGGGGAGCAGCACAAGTTTACCGCCCAACGCTTGGC 240
Qy 241 ACTAGGATCCAGAATGGCTACAACTGCTCCCTGATGGTGGCGCAATGGCTGAAATCCAA 300
```

```
Db 241 ACTAGGATCCAGAATGGCTACAACTGCTCCCTGATGGTGGCGCAATGGCTGAAATCCAA 300
Qy 301 GTACTACAGACTTTGTGATAAGCGTGAAGCTTGGGGCATCGTCTAGAAAACGCTGGCCAC 360
Db 301 GTACTACAGACTTTGTGATAAGCGTGAAGCTTGGGGCATCGTCTAGAAAACGCTGGCCAC 360
Qy 361 AGCGGGGTTGTGACCTCGGTGGCCTTCAATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420
Db 361 AGCGGGGTTGTGACCTCGGTGGCCTTCAATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420
Qy 421 GCAGGACTCCAA CAGCGGAAAAATGCTCCTACTCAAGTTTCTTCTCTCTCTGGGTGTGTT 480
Db 421 GCAGGACTCCAA CAGCGGAAAAATGCTCCTACTCAAGTTTCTTCTCTCTGGGTGTGTT 480
Qy 481 GGGCATCTTTGGCCTCACCTTCGCTTCAATCGGACTGGAGCGGAGCAGAGGGCCAC 540
Db 481 GGGCATCTTTGGCCTCACCTTCGCTTCAATCGGACTGGAGCGGAGCAGAGGGCCAC 540
Qy 541 AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Db 541 AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Qy 601 TGTCACTGTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTTCCCTGTTGGTGAATTCGGG 660
Db 601 TGTCACTGTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTTCCCTGTTGGTGAATTCGGG 660
Qy 661 TCTGGCGTGGGCTTCAGCCTAGTCCAGGATGTTATCGCTATTTGAATATATTTGCTCTGAC 720
Db 661 TCTGGCGTGGGCTTCAGCCTAGTCCAGGATGTTATCGCTATTTGAATATATTTGCTCTGAC 720
Qy 721 CATGAATAGGACCAACGTCAATGTCTTTTCTGAGCTTTCCGCTCTCTCGCAATGAAGA 780
Db 721 CATGAATAGGACCAACGTCAATGTCTTTTCTGAGCTTTCCGCTCTCTCGCAATGAAGA 780
Qy 781 CTTTGTCTCTCTGCTCACCTACGTCTCTTCTTGTGATGGCGCTGACCTTCTCTCATGCTC 840
Db 781 CTTTGTCTCTCTGCTCACCTACGTCTCTTCTTGTGATGGCGCTGACCTTCTCTCATGCTC 840
Qy 841 CTTACCTTCTGTGGTTCCTTTCACGGGCTGGAAGAGACATGGGGCCCACTCTACCTCAC 900
Db 841 CTTACCTTCTGTGGTTCCTTTCACGGGCTGGAAGAGACATGGGGCCCACTCTACCTCAC 900
Qy 901 GATGCTCTCTCCATTGGCCATCTGGGTGGCTGGATCACCTGCTCATGCTCTCTGACTT 960
Db 901 GATGCTCTCTCCATTGGCCATCTGGGTGGCTGGATCACCTGCTCATGCTCTCTGACTT 960
Qy 961 TGAACCGAGTGGGATGACACCATCTCAGCTCCGCTTTGGCTGCCAATGGCTGGGTGTT 1020
Db 961 TGAACCGAGTGGGATGACACCATCTCAGCTCCGCTTTGGCTGCCAATGGCTGGGTGTT 1020
Qy 1021 CCTGTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGACACGAAACCCCATGGA 1080
Db 1021 CCTGTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGACACGAAACCCCATGGA 1080
Qy 1081 TTATCTCTGTTGAGGATGCTTTCTGTAACCTCAA CTGTTGAAGAGAGCTATGGTGTGGA 1140
Db 1081 TTATCTCTGTTGAGGATGCTTTCTGTAACCTCAA CTGTTGAAGAGAGCTATGGTGTGGA 1140
Qy 1141 GAACGAGGCTACTCTCAAGAGGAAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCT 1200
Db 1141 GAACGAGGCTACTCTCAAGAGGAAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCT 1200
Qy 1201 CTATGCCCTTATTCACACATTTTCAGCTGCGAACCAGCTCCCCAAAAGAAATTCCTC 1260
Db 1201 CTATGCCCTTATTCACACATTTTCAGCTGCGAACCAGCTCCCCAAAAGAAATTCCTC 1260
Qy 1261 CATCCCAAGGGCCACGCTTGGCGAGCCTTACAAAGACTATGAAGTAAAGAAAGAGGG 1320
Db 1261 CATCCCAAGGGCCACGCTTGGCGAGCCTTACAAAGACTATGAAGTAAAGAAAGAGGG 1320
Qy 1321 CAGCTAACTCTGTCTGAAGAGTGGGACAAAATGACGCGGGCGGAGATCTTAGCGGAGC 1380
Db 1321 CAGCTAACTCTGTCTGAAGAGTGGGACAAAATGACGCGGGCGGAGATCTTAGCGGAGC 1380
```

Db 1321 CAGCTAACTCTGCTCTGAAGAGTGGGCAAAATGACGCCGGCGGCGAGATCTAGCGGGAGC 1380
Qy 1381 TCAAAAGGATGTGGGCAAAATCTGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGA 1440
Db 1381 TCAAAAGGATGTGGGCAAAATCTGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGA 1440
Qy 1441 ACAGTTTGCTCCCTCCCGACGCTCAACACAATTCCTTCATGCTGGGGCTGATGTGGGT 1500
Db 1441 ACAGTTTGCTCCCTCCCGACGCTCAACACAATTCCTTCATGCTGGGGCTGATGTGGGT 1500
Qy 1501 AGTAAGACTCCAGTCTTGAAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTTTGG 1560
Db 1501 AGTAAGACTCCAGTCTTGAAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTTTGG 1560
Qy 1561 ATACTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTTCAGCCCTTACTCTTTTGT 1620
Db 1561 ATACTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTTCAGCCCTTACTCTTTTGT 1620
Qy 1621 GTTTTTGAAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGTGCGATCACAG 1680
Db 1621 GTTTTTGAAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGTGCGATCACAG 1680
Qy 1681 CCCAGTGACGCTCGACACCTGTGCTCAAGCAATCTCCCATCTCCCAAGTG 1740
Db 1681 CCCAGTGACGCTCGACACCTGTGCTCAAGCAATCTCCCATCTCCCAAGTG 1740
Qy 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTGTGTTATTT 1800
Db 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTGTGTTATTT 1800
Qy 1801 CCATGGAATAAAGGTCTGTGCTCTGAGCTCAGCTCAGCTGAGTGCAGTGTAGGGGCTG 1860
Db 1801 CCATGGAATAAAGGTCTGTGCTCTGAGCTCAGCTCAGCTGAGTGCAGTGTAGGGGCTG 1860
Qy 1861 CTCCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
Db 1861 CTCCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
Qy 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCCAAC 1980
Db 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCCAAC 1980
Qy 1981 TTGCTGGACCCCGCTCCCAACCTCTTGGCTGGGTAGGAGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGACCCCGCTCCCAACCTCTTGGCTGGGTAGGAGGCTAAAGATCACCC 2040
Qy 2041 TAAATTTACTCATCTCTCTAGTGCTCCTCAATTTGGGCTCAGCAGCTCCCGACCA 2100
Db 2041 TAAATTTACTCATCTCTCTAGTGCTCCTCAATTTGGGCTCAGCAGCTCCCGACCA 2100
Qy 2101 ATTCAAGTCAACCCTCTCTTTCTTGCACTGTCCCAAACTTGTGTCAATTTCCAGATC 2160
Db 2101 ATTCAAGTCAACCCTCTCTTTCTTGCACTGTCCCAAACTTGTGTCAATTTCCAGATC 2160
Qy 2161 TAATCTCCCTACGCTCGCCAGGAATTTCTTCAGACCTCATAGCACAGCCGGTTG 2220
Db 2161 TAATCTCCCTACGCTCGCCAGGAATTTCTTCAGACCTCATAGCACAGCCGGTTG 2220
Qy 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTCAAATTCCTGGGCTGATCTCT 2280
Db 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTCAAATTCCTGGGCTGATCTCT 2280
Qy 2281 CTCATCTTGCACCCCCAACCTCTGTAAATAGATTTTACCGCATTTACGGCTGCAATCTGTAA 2340
Db 2281 CTCATCTTGCACCCCCAACCTCTGTAAATAGATTTTACCGCATTTACGGCTGCAATCTGTAA 2340
Qy 2341 GTGGGATGGTCTCCTAAATGGAGGAGTGTTCATTTGTAATAAGTATTTTCACTGAGTAT 2400
Db 2341 GTGGGATGGTCTCCTAAATGGAGGAGTGTTCATTTGTAATAAGTATTTTCACTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGGTGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456

RESULT 6
US-10-775-920-10
; Sequence 10, Application US/10775920
; Publication No. US2004017574A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 00108
; CURRENT APPLICATION NUMBER: US/10/775,920
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-10

Query Match 100.0%; Score 2456; DB 9; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATACAGCATGAAGTGCCTGGAACTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC 60
Db 1 ATACAGCATGAAGTGCCTGGAACTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC 60
Qy 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGCGCCCTTTATA 120
Db 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGCGCCCTTTATA 120
Qy 121 ACAACTGCTCAGATGCGAGGCGGGATAGTGTCAAGGTCTCCCGAGCAGCTGAGGAG 180
Db 121 ACAACTGCTCAGATGCGAGGCGGGATAGTGTCCAAGGTCTCCCGAGCAGCTGAGGAG 180
Qy 181 CTCGCTGCTGCCCTCTTTCGCGCGGGAAGCAGCAGCAAGTTTCAGCGCCAAAGCCTTGGC 240
Db 181 CTCGCTGCTGCCCTCTTTCGCGCGGGAAGCAGCAGCAAGTTTCAGCGCCAAAGCCTTGGC 240
Qy 241 ACTAGGGTCCAGAATGGCTTACAAAGTCCCTGATGGTTGCGCAATGCGCTGAAATCCAA 300
Db 241 ACTAGGGTCCAGAATGGCTTACAAAGTCCCTGATGGTTGCGCAATGCGCTGAAATCCAA 300
Qy 301 GTACTACAGACTTTGTGATAGGCTGAAGCTTTGGGGCATCGTCTAGAAACGGTGGCCAC 360
Db 301 GTACTACAGACTTTGTGATAGGCTGAAGCTTTGGGGCATCGTCTAGAAACGGTGGCCAC 360
Qy 361 AGCCGGGTGTGACCTCGGTGGCCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGT 420
Db 361 AGCCGGGTGTGACCTCGGTGGCCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGT 420
Qy 421 GCAGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTGGGTGTGT 480
Db 421 GCAGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTGGGTGTGT 480
Qy 481 GGGCATCTTTGGCCCTCACCTTCGCTTCAATTCATCGAGCTGGAGCGGAGCACAGGGCCAC 540
Db 481 GGGCATCTTTGGCCCTCACCTTCGCTTCAATTCATCGAGCTGGAGCGGAGCACAGGGCCAC 540
Qy 541 AGCCTTCTTCTCTTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTCTGCTCATGC 600
Db 541 AGCCTTCTTCTCTTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTCTGCTCATGC 600
Qy 601 TGTCACTGACAAAGTCTGTCGGGGAGGAGCCCTTTTCCCTGTTGGTGAATTCGGG 660
Db 601 TGTCACTGACAAAGTCTGTCGGGGAGGAGCCCTTTTCCCTGTTGGTGAATTCGGG 660
Qy 661 TCTGGCCGTGGGCTTCAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGCTCTGAC 720

Db 661 TCTGGCGGTGGGCTTTCAGGCTAGTCCAGGATGTTATCGCTATTGAATATATGTCCTGAC 720
Qy 721 CATGAATAGAACCAACGTCATGTCCTTTCTGAGCTTTCCGCTCCGTCGCAATGAAGA 780
Db 721 CATGAATAGAACCAACGTCATGTCCTTTCTGAGCTTTCCGCTCCGTCGCAATGAAGA 780
Qy 781 CTTTGTCTCTCTGCTCAGCTACGTCCTCTCTTGTGAGGGCTGACCTTCTCTCATGTCCTC 840
Db 781 CTTTGTCTCTCTGCTCAGCTACGTCCTCTCTTGTGAGGGCTGACCTTCTCTCATGTCCTC 840
Qy 841 CTTTCACTTCTGTCGTTTCTTCAAGGGCTGGAAGAGACATGAGGGCCACATCTACCTCAC 900
Db 841 CTTTCACTTCTGTCGTTTCTTCAAGGGCTGGAAGAGACATGAGGGCCACATCTACCTCAC 900
Qy 901 GATGCTCTCTCATTGCCATCTGGGTGGCTGGATCACCCCTGCTCATGCTTCTGACTT 960
Db 901 GATGCTCTCTCATTGCCATCTGGGTGGCTGGATCACCCCTGCTCATGCTTCTGACTT 960
Qy 961 TGACCCGAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT 1020
Db 961 TGACCCGAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGTT 1020
Qy 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACCAACCCCATGGA 1080
Db 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACCAACCCCATGGA 1080
Qy 1081 TTATCTGTTAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGACTATGGTGTGA 1140
Db 1081 TTATCTGTTAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGACTATGGTGTGA 1140
Qy 1141 GAACAGAGCTACTCTCAAGAGAAATCACTCAAGCTTTTGAAGACAGGGGACACGCT 1200
Db 1141 GAACAGAGCTACTCTCAAGAGAAATCACTCAAGCTTTTGAAGACAGGGGACACGCT 1200
Qy 1201 CTATGCCCCCTATTTCACACATTTTCAGCTGAGAACCAAGCTTCCCAAGAAAGAAATTC 1260
Db 1201 CTATGCCCCCTATTTCACACATTTTCAGCTGAGAACCAAGCTTCCCAAGAAAGAAATTC 1260
Qy 1261 CATCCACGGGCCCAACGCTTGGCGGAGCCCTTCAAGAGACTATGAAGTAAAGAAAGAGG 1320
Db 1261 CATCCACGGGCCCAACGCTTGGCGGAGCCCTTCAAGAGACTATGAAGTAAAGAAAGAGG 1320
Qy 1321 CAGCTAACTCTGCTGAGAGTGGGACAAATCAGCCGGCGGAGACTACGGGAGC 1380
Db 1321 CAGCTAACTCTGCTGAGAGTGGGACAAATCAGCCGGCGGAGACTACGGGAGC 1380
Qy 1381 TCAAGGGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
Db 1381 TCAAGGGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
Qy 1441 ACAGTTTGGCTCCCTCCAGGCTCAACCAATTTCTTCATGCTGGGGCTGATGGGCT 1500
Db 1441 ACAGTTTGGCTCCCTCCAGGCTCAACCAATTTCTTCATGCTGGGGCTGATGGGCT 1500
Qy 1501 AGTAAGACTCCAGTCTTGAAGGGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
Db 1501 AGTAAGACTCCAGTCTTGAAGGGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
Qy 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTGAAGCCCTTACTCTTTTGT 1620
Db 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTGAAGCCCTTACTCTTTTGT 1620
Qy 1621 GTTTTTTGAACAGGATCTTGTCTCAGCCAGGCTTGAAGTGGGATCAGACAG 1680
Db 1621 GTTTTTTGAACAGGATCTTGTCTCAGCCAGGCTTGAAGTGGGATCAGACAG 1680
Qy 1681 CCAGTGGACCTCGACCACTGCTCAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db 1681 CCAGTGGACCTCGACCACTGCTCAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Qy 1741 CTGGGATGACGGCGTGAGCCACAGCTCCAGGCTAGGCCCTTAACTTCTGCTGTTATTT 1800

Db 1741 CTGGGATGACGGCGTGAGCCACAGCTCCAGCCTAGGCCCTTAACTCTGCTGTTATTT 1800
Qy 1801 CANTGACTAAAGGTCTGTCATCTGAGCTCAAGCTGGCTCACAAGCTCTAGGGGCTG 1860
Db 1801 CANTGACTAAAGGTCTGTCATCTGAGCTCAAGCTGGCTCACAAGCTCTAGGGGCTG 1860
Qy 1861 CTCCTCTAACTACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACTGTCATATCT 1920
Db 1861 CTCCTCTAACTACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACTGTCATATCT 1920
Qy 1921 GAGCAAAAATAGCAAAAGGCTCTCTCAGCCCACTGGCCTGAAATCTACACTGGAAGCAAC 1980
Db 1921 GAGCAAAAATAGCAAAAGGCTCTCTCAGCCCACTGGCCTGAAATCTACACTGGAAGCAAC 1980
Qy 1981 TTGCTGGCACCCCGCTCCCAACCTTCTTGGCTGGTAGGAGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCCGCTCCCAACCTTCTTGGCTGGTAGGAGGCTAAAGATCACCC 2040
Qy 2041 TAAATTTACTCATCTCTCTAGTGCCTCAGCTCAGCTTGGGCTCAGAGCTCCCAAGCA 2100
Db 2041 TAAATTTACTCATCTCTCTAGTGCCTCAGCTCAGCTTGGGCTCAGAGCTCCCAAGCA 2100
Qy 2101 ATTACAGGTGACCCCTCTCTTCTGCACTGTCCCAAACTTGTCTGCAATTCGAGATC 2160
Db 2101 ATTACAGGTGACCCCTCTCTTCTGCACTGTCCCAAACTTGTCTGCAATTCGAGATC 2160
Qy 2161 TAACTCCCGCTAGGCTCTGCGAGGAATTTCTTTCAGACCTCAGTACAGCAAGCCGGTTG 2220
Db 2161 TAACTCCCGCTAGGCTCTGCGAGGAATTTCTTTCAGACCTCAGTACAGCAAGCCGGTTG 2220
Qy 2221 CTCCTTGTGAGGAAATTTGTAGTCACTTCTCACTTCAAAATCTTGGGCTGATCTTCT 2280
Db 2221 CTCCTTGTGAGGAAATTTGTAGTCACTTCTCACTTCAAAATCTTGGGCTGATCTTCT 2280
Qy 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTTACCGATTTACGGCTGCAATTCGTAA 2340
Db 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTTACCGATTTACGGCTGCAATTCGTAA 2340
Qy 2341 GTGGGATGGTCTCTTAATGGAGGAGTTCATTGTATATAAGTTATTACCTGAGTAT 2400
Db 2341 GTGGGATGGTCTCTTAATGGAGGAGTTCATTGTATATAAGTTATTACCTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGGTGGCCTCTTTCATGTTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGGCCTCTTTCATGTTGGTGGCAGCAAAAAA 2456

RESULT 7

US-10-936-626-40
; Sequence 40, Application US/10936626
; Publication No. US20050106644A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21

; PRIOR APPLICATION NUMBER: US 10/241,220									
; PRIOR FILING DATE: 2002-09-11									
; PRIOR APPLICATION NUMBER: US 10/177,488									
; PRIOR FILING DATE: 2002-06-19									
; PRIOR APPLICATION NUMBER: US 60/299,500									
; PRIOR FILING DATE: 2001-06-20									
; PRIOR APPLICATION NUMBER: US 60/301,880									
; PRIOR FILING DATE: 2001-06-29									
; PRIOR APPLICATION NUMBER: US 60/323,268									
; PRIOR FILING DATE: 2001-09-18									
; PRIOR APPLICATION NUMBER: US 60/557,116									
; PRIOR FILING DATE: 2004-03-26									
; PRIOR APPLICATION NUMBER: US 60/598,899									
; PRIOR FILING DATE: 2004-08-04									
; NUMBER OF SEQ ID NOS: 154									
; SEQ ID NO 40									
; LENGTH: 2456									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-936-626-40									
Query Match									
Best Local Similarity 100.0%; Score 2456; DB 10; Length 2456;									
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATAACAGCATGAAGTGCCTGGAACTGGAAATAGGGGTGTCTCTCCCTCGACCCCTCCCC	60						
DB	1	ATAACAGCATGAAGTGCCTGGAACTGGAAATAGGGGTGTCTCTCCCTCGACCCCTCCCC	60						
QY	61	TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCCTCCGGGAGGGCCGCTTTATA	120						
DB	61	TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCCTCCGGGAGGGCCGCTTTATA	120						
QY	121	ACAACTGCTCAGATGCGAGGCGGGATAGCTGTCCAAAGTCTCCGCCAGCACTGAGGAG	180						
DB	121	ACAACTGCTCAGATGCGAGGCGGGATAGCTGTCCAAAGTCTCCGCCAGCACTGAGGAG	180						
QY	181	CTCGCTGTGCTCCCTCTTCGCGCGGGGAAAGCAGCACCAGTTTACGGCCAAACGCTTTGGC	240						
DB	181	CTCGCTGTGCTCCCTCTTCGCGCGGGGAAAGCAGCACCAGTTTACGGCCAAACGCTTTGGC	240						
QY	241	ACTAGGGTCCAGATGGCTTACACAGTCCCTCATGGTTGCGCGCAATGGCTGAATCCAA	300						
DB	241	ACTAGGGTCCAGATGGCTTACACAGTCCCTCATGGTTGCGCGCAATGGCTGAATCCAA	300						
QY	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTTGGGCGATCGTCTAGAAAACGGTGGCCAC	360						
DB	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTTGGGCGATCGTCTAGAAAACGGTGGCCAC	360						
QY	361	AGCCGGGGTTGACCTCGGTGGCCCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420						
DB	361	AGCCGGGGTTGACCTCGGTGGCCCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420						
QY	421	GCAGGACTCCAGAGCGGAAAATGCTGCTACTCATGTTTCTTCTTCTCTCGGTGTGT	480						
DB	421	GCAGGACTCCAGAGCGGAAAATGCTGCTACTCATGTTTCTTCTTCTCTCGGTGTGT	480						
QY	481	GGGCACTTTTGGCTCACCTCGCTTCATCATCGGACTTGGACGGGAGCACAGGGCCCAAC	540						
DB	481	GGGCACTTTTGGCTCACCTCGCTTCATCATCGGACTTGGACGGGAGCACAGGGCCCAAC	540						
QY	541	ACGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTGCTATGC	600						
DB	541	ACGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTGCTATGC	600						
QY	601	TGTCAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTCCCTGTTGCTGATTTCTGGG	660						
DB	601	TGTCAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTCCCTGTTGCTGATTTCTGGG	660						
QY	661	TCTGGCCGTGGGCTTACAGCTTAGTCAGGATGTTATCGCTATTGAATATATTGCTCTGAC	720						
DB	661	TCTGGCCGTGGGCTTACAGCTTAGTCAGGATGTTATCGCTATTGAATATATTGCTCTGAC	720						

QY	721	CATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCCGCTCTCTGTCGCAATGAAGA	780
DB	721	CATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCCGCTCTCTGTCGCAATGAAGA	780
QY	781	CTTTGTCTCTCTGCTCACTACGTCCTCTCTTTGATGGCGTGACCTTCTCATGTCTCTC	840
DB	781	CTTTGTCTCTCTGCTCACTACGTCCTCTCTTTGATGGCGTGACCTTCTCATGTCTCTC	840
QY	841	CTTACACCTTCTGTGGTTCTCTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTAC	900
DB	841	CTTACACCTTCTGTGGTTCTCTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTAC	900
QY	901	GATGCTCTCTCCATTGGCATCTGGGTGSCCTTGGATCAACCTTCTCATGCTTCTCTGACTT	960
DB	901	GATGCTCTCTCCATTGGCATCTGGGTGSCCTTGGATCAACCTTCTCATGCTTCTCTGACTT	960
QY	961	TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCTGCAATGGCTGGGTGTT	1020
DB	961	TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCTGCAATGGCTGGGTGTT	1020
QY	1021	CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGTCTCAAAAGCAACGAAACCCCATGGA	1080
DB	1021	CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGTCTCAAAAGCAACGAAACCCCATGGA	1080
QY	1081	TTATCCTGTTGAGGATGCTTTCTGTAAACCTCAACTCTGTTGAAGAGAGCTATGGTGTGA	1140
DB	1081	TTATCCTGTTGAGGATGCTTTCTGTAAACCTCAACTCTGTTGAAGAGAGCTATGGTGTGA	1140
QY	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGTTTGAAGAGACAGGGGACACGCT	1200
DB	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGTTTGAAGAGACAGGGGACACGCT	1200
QY	1201	CTATGCCCTTATTTCCACACATTTTTCAGCTGAGAACCAAGCTTCCCAAAAGAAATTTCT	1260
DB	1201	CTATGCCCTTATTTCCACACATTTTTCAGCTGAGAACCAAGCTTCCCAAAAGAAATTTCT	1260
QY	1261	CATCCCAACGGGCGCAGCTTGGCCAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG	1320
DB	1261	CATCCCAACGGGCGCAGCTTGGCCAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG	1320
QY	1321	CAGCTTAACCTCTCTCTGAAGAGTGGGACAAATGACGCGGGCGGACAGATCTAGCGGGAGC	1380
DB	1321	CAGCTTAACCTCTCTCTGAAGAGTGGGACAAATGACGCGGGCGGACAGATCTAGCGGGAGC	1380
QY	1381	TCAAAGGGGATGGGCGAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA	1440
DB	1381	TCAAAGGGGATGGGCGAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA	1440
QY	1441	ACAGTTTGGCTCTCCCTCCAGCCCTCAACCAATTTCTTCCATGCTGGGGCTGATGGGCT	1500
DB	1441	ACAGTTTGGCTCTCCCTCCAGCCCTCAACCAATTTCTTCCATGCTGGGGCTGATGGGCT	1500
QY	1501	AGTAAAGACTCCAGTTCTTTAGAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTCTGG	1560
DB	1501	AGTAAAGACTCCAGTTCTTTAGAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTCTGG	1560
QY	1561	ATACTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTGTAGACCCCTTACTCTTTTGT	1620
DB	1561	ATACTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTGTAGACCCCTTACTCTTTTGT	1620
QY	1621	GTTTTTTGAACAGAGATCTTGTCTGTCAACGAGCTTGAAGTGCAGTGGCGATCACAG	1680
DB	1621	GTTTTTTGAACAGAGATCTTGTCTGTCAACGAGCTTGAAGTGCAGTGGCGATCACAG	1680
QY	1681	CCAGTGCAGCTCGAACCACTGTGCTCAAGCAATCCTCCCATCTCCATCTCCCAAGTG	1740
DB	1681	CCAGTGCAGCTCGAACCACTGTGCTCAAGCAATCCTCCCATCTCCATCTCCCAAGTG	1740
QY	1741	CTGGGATGACAGGCGTAGCCACAGCTCCAGGCTAGGCGCTTAATCTTGTCTGTTATTTT	1800
DB	1741	CTGGGATGACAGGCGTAGCCACAGCTCCAGGCTAGGCGCTTAATCTTGTCTGTTATTTT	1800
QY	1801	CCATGGACTAAAGGTCTGGTCTATCTGAGCTCACGCTGAGCTCACAGCTCTAGGGGCTG	1860

QY 721 CATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTFCGCAATGAAGA 780
DB 721 CATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTFCGCAATGAAGA 780
QY 781 CTTTCTGCTCCTGCTCAACCTAGCTCCTCTCTCTTGTATGGCGCTGACCTTCTCATGCTCCTC 840
DB 781 CTTTGTGCTCCTGCTCAACCTAGCTCCTCTCTCTTGTATGGCGCTGACCTTCTCATGCTCCTC 840
QY 841 CTTTCACTTCTGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 841 CTTTCACTTCTGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 GATGCTCCTCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 GATGCTCCTCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CTTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 CTTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TTATCCTGTTGAGGATGCTTTCTGTAACCTCAACTGCTGTAAGAGAGAGCTATGTTGTTGGA 1140
DB 1081 TTATCCTGTTGAGGATGCTTTCTGTAACCTCAACTGCTGTAAGAGAGAGCTATGTTGTTGGA 1140
QY 1141 GAACGAGCCTACTCTCAAGAGGAAATCACTCAAGTCTTGAAGAGAGAGGAGGACACGCT 1200
DB 1141 GAACGAGCCTACTCTCAAGAGGAAATCACTCAAGTCTTGAAGAGAGAGGAGGACACGCT 1200
QY 1201 CTATGCCCCCTATTCACACATTTTTCAGCTGAGAACACGAGCTCCGCCAAAGAAATTTCTC 1260
DB 1201 CTATGCCCCCTATTCACACATTTTTCAGCTGAGAACACGAGCTCCGCCAAAGAAATTTCTC 1260
QY 1261 CATCCCAAGGGCCACGCTTGGCCGAGCCCTTACAAAGCTATGAAAGTAAAGAAAGAGGG 1320
DB 1261 CATCCCAAGGGCCACGCTTGGCCGAGCCCTTACAAAGCTATGAAAGTAAAGAAAGAGGG 1320
QY 1321 CAGCTAACTCTCTCTGAAGAGTGGCAAAATGACGCGGGCGGAGATCTAGCGGGAGC 1380
DB 1321 CAGCTAACTCTCTCTGAAGAGTGGCAAAATGACGCGGGCGGAGATCTAGCGGGAGC 1380
QY 1381 TCAAAAGGGATGGGCGAAATCTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
DB 1381 TCAAAAGGGATGGGCGAAATCTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
QY 1441 ACAGTTTGGCTCCCTCCAGCCTCAACCAATTTCTCCATGCTGGGGCTGATGGGCT 1500
DB 1441 ACAGTTTGGCTCCCTCCAGCCTCAACCAATTTCTCCATGCTGGGGCTGATGGGCT 1500
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGAGTATTTTTTTTTTTTGTCTCATCCTTTGG 1560
DB 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGAGTATTTTTTTTTTTTGTCTCATCCTTTGG 1560
QY 1561 ATACTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTATAGACCCCTTACTCTTTTGTGTT 1620
DB 1561 ATACTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTATAGACCCCTTACTCTTTTGTGTT 1620
QY 1621 GTTTTTTGAACAGAGTCTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1621 GTTTTTTGAACAGAGTCTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
QY 1681 CCCAGTGCAGCTGCACACCTGCTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGG 1740
DB 1681 CCCAGTGCAGCTGCACACCTGCTGCTCAAGCAATCTCCCATCTCCCAAGG 1740
QY 1741 CTGGGATGACAGGCGGTGAGCCACAGCTCCAGCCCTAGGCGCTTAAATCTTGTGCTGTTATTTT 1800
DB 1741 CTGGGATGACAGGCGGTGAGCCACAGCTCCAGCCCTAGGCGCTTAAATCTTGTGCTGTTATTTT 1800
QY 1801 CCATGACTAAAGGCTTGTGCTATCTGAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

DB 1801 CCATGACTAAAGGCTTGTGCTATCTGAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
QY 1861 CTTCTCTTAACCTACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGTCATATCT 1920
DB 1861 CTTCTCTTAACCTACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGTCATATCT 1920
QY 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCTGAAATCTACACTGGAAGCAAC 1980
DB 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCTGAAATCTACACTGGAAGCAAC 1980
QY 1981 TTGCTGGCACCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
DB 1981 TTGCTGGCACCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY 2041 TAAATTTACTCATCTCTCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 2041 TAAATTTACTCATCTCTCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2101 ATTCAAGGTACCTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 2101 ATTCAAGGTACCTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
QY 2161 TAAATCTCCCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
DB 2161 TAAATCTCCCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY 2221 CTTCTTGTGAGGAGAAATTTGATGATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
DB 2221 CTTCTTGTGAGGAGAAATTTGATGATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY 2281 CTTATCTTGCACCCCAACCTCTGTAATAGATTTACCGCATTTACGGCTGCTGCTGCTGCTGCTGCTGCT 2340
DB 2281 CTTATCTTGCACCCCAACCTCTGTAATAGATTTACCGCATTTACGGCTGCTGCTGCTGCTGCTGCTGCT 2340
QY 2341 GTGGGCACTGCTCTCTTAATGAGGAGTGTTCATTTATTAATAGTTTATTCACCTGAGTAT 2400
DB 2341 GTGGGCACTGCTCTCTTAATGAGGAGTGTTCATTTATTAATAGTTTATTCACCTGAGTAT 2400
QY 2401 GCAATAAAGATGTGCTGGCCACTCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2456
DB 2401 GCAATAAAGATGTGCTGGCCACTCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2456

RESULT 9

US-11-169-041-32
; Sequence 32, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-169-041-32

Query Match 100.0%; Score 2456; DB 16; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAACGACATGAAGTGGCGTGAACCTGGAATAGGCGTGTCTCTCCCTCGACCTCCCC 60

QY 1021 CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
Db 1021 CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
QY 1081 TTATCCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGCTGTGGA 1140
Db 1081 TTATCCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGCTGTGGA 1140
QY 1141 GAACAGAGCTACTCTCAAGAGAAATCACTCAAGGTTTGAAGAGACAGGGGACAGCT 1200
Db 1141 GAACAGAGCTACTCTCAAGAGAAATCACTCAAGGTTTGAAGAGACAGGGGACAGCT 1200
QY 1201 CTATGCCCCCTATTCCACACATTTTCAGCTGCAGAACCCAGCTCCCCCAAGAAAGAAATTC 1260
Db 1201 CTATGCCCCCTATTCCACACATTTTCAGCTGCAGAACCCAGCTCCCCCAAGAAAGAAATTC 1260
QY 1261 CATCCCAACGGGCCACGCTTGGCCGAGCCCTTACAAGAGCTATGAAGTAAAGAAAGAGGG 1320
Db 1261 CATCCCAACGGGCCACGCTTGGCCGAGCCCTTACAAGAGCTATGAAGTAAAGAAAGAGGG 1320
QY 1321 CAGCTAACTCTGTCCTGAAGAGTGGGAACAAATGCAGCCGGGCGGAGATCTAGCGGGAGC 1380
Db 1321 CAGCTAACTCTGTCCTGAAGAGTGGGAACAAATGCAGCCGGGCGGAGATCTAGCGGGAGC 1380
QY 1381 TCAAGGGGATGCGGCGAAATCTTGAGTCTCTGAGAAACCTGTACAAGACACTACGGGA 1440
Db 1381 TCAAGGGGATGCGGCGAAATCTTGAGTCTCTGAGAAACCTGTACAAGACACTACGGGA 1440
QY 1441 ACAGTTTGCTCCCTCCAGGCTCAACCAACATTTCTTCCATGCTGGGCTGATGCGGCT 1500
Db 1441 ACAGTTTGCTCCCTCCAGGCTCAACCAACATTTCTTCCATGCTGGGCTGATGCGGCT 1500
QY 1501 AGTAAGACTCCAGTCTTAGAGCGCTGATGATTTTTTTTTTTTTTTTGTCTCATCTTTGG 1560
Db 1501 AGTAAGACTCCAGTCTTAGAGCGCTGATGATTTTTTTTTTTTTTTTGTCTCATCTTTGG 1560
QY 1561 ATACTTCTTTTAACTGGGAGTCTCAGGCAACTCAAGTTTGTAGACCTTACTCTTTTGT 1620
Db 1561 ATACTTCTTTTAACTGGGAGTCTCAGGCAACTCAAGTTTGTAGACCTTACTCTTTTGT 1620
QY 1621 GTTTTTTGAACAGGATCTTGCTCTGTCAACAGGCTTGAAGTGCAGTGGGATCACAG 1680
Db 1621 GTTTTTTGAACAGGATCTTGCTCTGTCAACAGGCTTGAAGTGCAGTGGGATCACAG 1680
QY 1681 CCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db 1681 CCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
QY 1741 CTGGGATGACAGGCGTGAAGCACTCCAGGCTAGGCGCTTAATCTTGTCTGTTATTTT 1800
Db 1741 CTGGGATGACAGGCGTGAAGCACTCCAGGCTAGGCGCTTAATCTTGTCTGTTATTTT 1800
QY 1801 CCATGACTAAAGGCTTGCTGCTCAGCTCAGCTGCTGCTCAACAGCTCTAGGGGCTG 1860
Db 1801 CCATGACTAAAGGCTTGCTGCTCAGCTCAGCTGCTGCTCAACAGCTCTAGGGGCTG 1860
QY 1861 CTCCTCTAATCACTAGTGGGTTTGTGAGGCTCTGTGGGCCAGAGCAGACCTGCATATCT 1920
Db 1861 CTCCTCTAATCACTAGTGGGTTTGTGAGGCTCTGTGGGCCAGAGCAGACCTGCATATCT 1920
QY 1921 GAGCAAAATAGCAAAAGCTCTCTCAGGCCACTGGGCTGAACTCTACCTGGAAGCCAAC 1980
Db 1921 GAGCAAAATAGCAAAAGCTCTCTCAGGCCACTGGGCTGAACTCTACCTGGAAGCCAAC 1980
QY 1981 TTGCTGGCACCCCGCTCCCAACCCCTTTTGTGCTGGGTAGGAGAGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCCGCTCCCAACCCCTTTTGTGCTGGGTAGGAGAGGCTAAAGATCACCC 2040
QY 2041 TAAATTTACTCATCTCTAGTCTGCTCAATTTGGGCTCAGAGCTCCCGCAGCACCA 2100
Db 2041 TAAATTTACTCATCTCTAGTCTGCTCAATTTGGGCTCAGAGCTCCCGCAGCACCA 2100

QY 2101 ATTACAGGTACCCCTCTCTTCTTGTGCTGTCCTCCCAAACTTGTCTCAATTCCGAGATC 2160
Db 2101 ATTACAGGTACCCCTCTCTTCTTGTGCTGTCCTCCCAAACTTGTCTCAATTCCGAGATC 2160
QY 2161 TAATCTCCCTCCTACGCTCTGCCAGGAATTTCTTTAGACCTCTACTAGCACAGCCCGTTG 2220
Db 2161 TAATCTCCCTCCTACGCTCTGCCAGGAATTTCTTTAGACCTCTACTAGCACAGCCCGTTG 2220
QY 2221 CTCCTTGTGAGAGAAATTTGTAGATCAATTTCTCACTTCAAAATTCCTGGGCTGATCTCT 2280
Db 2221 CTCCTTGTGAGAGAAATTTGTAGATCAATTTCTCACTTCAAAATTCCTGGGCTGATCTCT 2280
QY 2281 CTCATCTTGCAACCCAACTCTGTAATAGATTTACCGATTTACGGCTGCAATTTCTGTAA 2340
Db 2281 CTCATCTTGCAACCCAACTCTGTAATAGATTTACCGATTTACGGCTGCAATTTCTGTAA 2340
QY 2341 GTGGGATGCTCTCTTAATGAGGAGTGTTCATTTGATTAATAGTTTATTCACCTGAGTAT 2400
Db 2341 GTGGGATGCTCTCTTAATGAGGAGTGTTCATTTGATTAATAGTTTATTCACCTGAGTAT 2400
QY 2401 GCAATAAGATGCTGGGCCACTCTTTCATGTTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAGATGCTGGGCCACTCTTTCATGTTGGTGGCAGCAAAAAA 2456

RESULT 11
US-10-198-846-10424
; Sequence 10424, Application US/10198846
; Publication No. US2003099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10424
; LENGTH: 4239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 4238, 4239
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10424

Query Match 99.5%; Score 2443; DB 6; Length 4239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATACAGCATGAAGTCCGTGGAACCTGGAATAGGCGTGTCTCTCCCTCGAACCTCCCC 60
Db 409 ATACAGCATGAAGTCCGTGGAACCTGGAATAGGCGTGTCTCTCCCTCGAACCTCCCC 468
QY 61 TCCTTGTCCCTCTGCTACCCCTCGCTCGTTCCCTCCCTCCGGCGAGGGCGGCTTTATA 120
Db 469 TCCTTGTCCCTCTGCTACCCCTCGCTCGTTCCCTCCCTCCGGCGAGGGCGGCTTTATA 528
QY 121 ACACTGCTCAGAGTGCAGAGGGCGGATAGCTGTCCAAGGTCTCCCGCAGCACTGAGGAG 180
Db 529 ACACTGCTCAGAGTGCAGAGGGCGGATAGCTGTCCAAGGTCTCCCGCAGCACTGAGGAG 588
QY 181 CTCGCTGCTGCCCTCTTTCGCGCGGGAAGAGCAACAAAGTTTCAAGGCAACGCTTTGGC 240
Db 589 CTCGCTGCTGCCCTCTTTCGCGCGGGAAGAGCAACAAAGTTTCAAGGCAACGCTTTGGC 648

QY	241	ACTAGGGTCCAGATGGCTCAACAGTCCCTGATGGTTGCCCAATGGGCTGAAATCCAA	300
Db	649	ACTAGGGTCCAGATGGCTTACACAGTCCCTGATGGTTGCCCAATGGCTGAAATCCAA	708
QY	301	GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGGCATCGCTAGAAACGGTGGCCAC	360
Db	709	GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGGCATCGCTAGAAACGGTGGCCAC	768
QY	361	AGCCGGGTTGTGACCTCGGTTGGCCCTTCAATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420
Db	769	AGCCGGGTTGTGACCTCGGTTGGCCCTTCAATGCTCACTCTCCGATCCTCGTCTGCAAGGT	828
QY	421	GCAGGACTCCAA CAGCGGAAAATGCTGCTACTCAGATTCTCTTCCCTCGGTTGTGT	480
Db	829	GCAGGACTCCAA CAGCGGAAAATGCTGCTACTCAGATTCTCTTCCCTCGGTTGTGT	888
QY	481	GGGCATCTTTGGGCTCACCTTGGCTTCCATCATCGGACTGGAGCGGAGCACAGGGCCAC	540
Db	889	GGGCATCTTTGGGCTCACCTTGGCTTCCATCATCGGACTGGAGCGGAGCACAGGGCCAC	948
QY	541	AGGCTTCTTCTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGGCTCATGC	600
Db	949	AGGCTTCTTCTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGGCTCATGC	1008
QY	601	TGTCAGTCTGACCAAGCTCGTCGGGGAGGAAGCCCTTTCCCTGTGGTGAATCTGGG	660
Db	1009	TGTCAGTCTGACCAAGCTCGTCGGGGAGGAAGCCCTTTCCCTGTGGTGAATCTGGG	1068
QY	661	TCGGGCTGGGCTTCCAGCTAGTCCAGGATGTTATCGCTATGATATATTTGCTCTGAC	720
Db	1069	TCGGGCTGGGCTTCCAGCTAGTCCAGGATGTTATCGCTATGATATATTTGCTCTGAC	1128
QY	721	CATGAATAGGACCAACGCTCAATGCTTTTCTGAGCTTTCCGCTCTCTCGCGCAATGAAGA	780
Db	1129	CATGAATAGGACCAACGCTCAATGCTTTTCTGAGCTTTCCGCTCTCTCGCGCAATGAAGA	1188
QY	781	CTTTGCTCTCTGCTCACTACCTGCTCTCTTTGATGGCGCTGACCTTCCCTCATGTCCTC	840
Db	1189	CTTTGCTCTCTGCTCACTACCTGCTCTCTTTGATGGCGCTGACCTTCCCTCATGTCCTC	1248
QY	841	CTTCACTTCTGTTGTTCTTCA CGGCTGGAGACATGGGCGCCACATCACTCAC	900
Db	1249	CTTCACTTCTGTTGTTCTTCA CGGCTGGAGACATGGGCGCCACATCACTCAC	1308
QY	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACCTGCTCATGCTTCTTGACTT	960
Db	1309	GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACCTGCTCATGCTTCTTGACTT	1368
QY	961	TGACCGCAGTGGGATGACCACTCTCAGCTCCGCTTTGGTGGCAATGGCTGGGTGTT	1020
Db	1369	TGACCGCAGTGGGATGACCACTCTCAGCTCCGCTTTGGTGGCAATGGCTGGGTGTT	1428
QY	1021	CCTGTGGCTTATGTAGTCCGATTTTGGCTGCTCACAAGCAACGAAACCCCATGGA	1080
Db	1429	CCTGTGGCTTATGTAGTCCGATTTTGGCTGCTCACAAGCAACGAAACCCCATGGA	1488
QY	1081	TTATCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTGTGGA	1140
Db	1489	TTATCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTGTGGA	1548
QY	1141	GAA CAGAGCTACTCTCAAGAGAAATCTCAAGGTTTTGAAGAGAGAGGAGACACGCT	1200
Db	1549	GAA CAGAGCTACTCTCAAGAGAAATCTCAAGGTTTTGAAGAGAGAGGAGACACGCT	1608
QY	1201	CTATGCCCTTATTCACACATTTTCAGCTGAGACACGCTCCCAAAAGAAATTTCTC	1260
Db	1609	CTATGCCCTTATTCACACATTTTCAGCTGAGACACGCTCCCAAAAGAAATTTCTC	1668
QY	1261	CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG	1320
Db	1669	CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG	1728

QY	1321	CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGCGAGCCGGCGGCAGATCTTAGCGGGAGC	1380
Db	1729	CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGCGAGCCGGCGGCAGATCTTAGCGGGAGC	1788
QY	1381	TCAAAAGGATGTGGCGGAAATCTTGAGTCTCTGAGAAAACCTGTACAACACTACGGGA	1440
Db	1789	TCAAAAGGATGTGGCGGAAATCTTGAGTCTCTGAGAAAACCTGTACAACACTACGGGA	1848
QY	1441	ACAGTTTGGCTCCCTCCAGCCCTCAACCAATTTCTTCCATGCTGGGGCTGATGGGCT	1500
Db	1849	ACAGTTTGGCTCCCTCCAGCCCTCAACCAATTTCTTCCATGCTGGGGCTGATGGGCT	1908
QY	1501	AGTAAGA CT CAGTCTTTAGAGCGCTGATGATTTTTTTTTTTTTTGTCTCATCTTTGG	1560
Db	1909	AGTAAGA CT CAGTCTTTAGAGCGCTGATGATTTTTTTTTTTTTTGTCTCATCTTTGG	1968
QY	1561	ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTT TAGACCTTACTCTTTTGT	1620
Db	1969	ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTT TAGACCTTACTCTTTTGT	2028
QY	1621	GTTTTTGAAACAGGATCTTGTCTGTCA CCGAGCTTGA GTGAGTGCATCACAG	1680
Db	2029	GTTTTTGAAACAGGATCTTGTCTGTCA CCGAGCTTGA GTGAGTGCATCACAG	2088
QY	1681	CCAGTGCAGCTCGACCACTGTGCTCAGCAATCTCTCCATCTCCCAAGTG	1740
Db	2089	CCAGTGCAGCTCGACCACTGTGCTCAGCAATCTCTCCATCTCCCAAGTG	2148
QY	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGCTTAGGCCCTTAATCTTGTGTTATTTT	1800
Db	2149	CTGGGATGACAGGCTGAGCCACAGCTCCAGCTTAGGCCCTTAATCTTGTGTTATTTT	2208
QY	1801	CCATGGA CTAAAGTCTGTGTCTCTGAGCTCAGCTCGTGGCTCA CACAGCTCTAGGGCC	1860
Db	2209	CCATGGA CTAAAGTCTGTGTCTCTGAGCTCAGCTCGTGGCTCA CACAGCTCTAGGGCC	2268
QY	1861	CTTCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGAGACCTTGCATATCT	1920
Db	2269	CTTCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGAGACCTTGCATATCT	2328
QY	1921	GAGCAAAATAGCAAAAGCTCTCTCAGCCACTGGCCCTGAATCTACACTGGAAGCAAC	1980
Db	2329	GAGCAAAATAGCAAAAGCTCTCTCAGCCACTGGCCCTGAATCTACACTGGAAGCAAC	2388
QY	1981	TTGTCTGCA CCCCCCTCCCAACCTTCTTGCTGGGTAGGAGAGCTTAAAGATCACCC	2040
Db	2389	TTGTCTGCA CCCCCCTCCCAACCTTCTTGCTGGGTAGGAGAGCTTAAAGATCACCC	2448
QY	2041	TAAATTTACTCATCTCTTAGTGTGCTCACAATTGGGCTCAGCAGCTCCCGACCA	2100
Db	2449	TAAATTTACTCATCTCTTAGTGTGCTCACAATTGGGCTCAGCAGCTCCCGACCA	2508
QY	2101	ATTACAGTCA CCCCCTCTTCTTCTTGCACTGCTCCCAAACTTGCTGTCAATCCGAGATC	2160
Db	2509	ATTACAGTCA CCCCCTCTTCTTCTTGCACTGCTCCCAAACTTGCTGTCAATCCGAGATC	2568
QY	2161	TAAATCTCCCTTACGCTCTGCGAGGAA TTTCTTTTACACCTCTAGCA CCAAGCCCGGTTG	2220
Db	2569	TAAATCTCCCTTACGCTCTGCGAGGAA TTTCTTTTACACCTCTAGCA CCAAGCCCGGTTG	2628
QY	2221	CTTCTTGTGAGGAA TTTGTAGATCATTTCTCAGTTCAAATTCCTGGGCTGATCTCT	2280
Db	2629	CTTCTTGTGAGGAA TTTGTAGATCATTTCTCAGTTCAAATTCCTGGGCTGATCTCT	2688
QY	2281	CTCATCTTGCACCCCAACCTCTGTAAATAGATTTTACCGCATTTAGGGCTGCA TTTGTAA	2340
Db	2689	CTCATCTTGCACCCCAACCTCTGTAAATAGATTTTACCGCATTTAGGGCTGCA TTTGTAA	2748
QY	2341	GTGGGATGCTCTCTTAATGAGGAGTGTTCATTTGTAATAATTAATTAATTAATTAATTAAT	2400
Db	2749	GTGGGATGCTCTCTTAATGAGGAGTGTTCATTTGTAATAATTAATTAATTAATTAATTAAT	2808
QY	2401	GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGCGAGCA	2443

QY 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGGCTAGGCCCTTAATCTTGCTGTTATTTT 1800
DB |||||
1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGGCTAGGCCCTTAATCTTGCTGTTATTTT 1800
QY 1801 CCATGAGCTAAAGGCTCTGCTCATCTGAGCTCAGCTGGCTCACACAGCTCTAGGGCCCTG 1860
DB |||||
1801 CCATGAGCTAAAGGCTCTGCTCATCTGAGCTCAGCTGGCTCACACAGCTCTAGGGCCCTG 1860
QY 1861 CTCTCTAACTCAAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
DB |||||
1861 CTCTCTAACTCAAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
QY 1921 GAGCAAAAATAGCAAAAGGCTCTCTCAGCCCACTGGCTGGAATCTACACTGGAAGCCAAAC 1980
DB |||||
1921 GAGCAAAAATAGCAAAAGGCTCTCTCAGCCCACTGGCTGGAATCTACACTGGAAGCCAAAC 1980
QY 1981 TTGCTGGCACCCCGCTCCCAACCTCTCTGCTGGGTAGGAGGCTAAAGATCACCC 2040
DB |||||
1981 TTGCTGGCACCCCGCTCCCAACCTCTCTGCTGGGTAGGAGGCTAAAGATCACCC 2040
QY 2041 TAAATTTACTCATCTCTCTAGTGTGCTCTCACAATTGGGCTCAGCAGCTCCCCAGCACCA 2100
DB |||||
2041 TAAATTTACTCATCTCTCTAGTGTGCTCTCACAATTGGGCTCAGCAGCTCCCCAGCACCA 2100
QY 2101 ATTACAGGTACCCCTCTCTTTCTTGCACTGTCCCAAACTTGCTGTCAAATCCGAGATC 2160
DB |||||
2101 ATTACAGGTACCCCTCTCTTTCTTGCACTGTCCCAAACTTGCTGTCAAATCCGAGATC 2160
QY 2161 TAAATCTCCCTTACGCTCTGCCAGGAATCTTTTCAGACCTCACTAGCAAGCCCGGTTG 2220
DB |||||
2161 TAAATCTCCCTTACGCTCTGCCAGGAATCTTTTCAGACCTCACTAGCAAGCCCGGTTG 2220
QY 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTTCAAATTCCTGGGCTGATATCT 2280
DB |||||
2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTTCAAATTCCTGGGCTGATATCT 2280
QY 2281 CTCATCTTGCAACCCCACTCTGTAAATAGATTTACCGCATTTACGGCTGCAATCTGTAA 2340
DB |||||
2281 CTCATCTTGCAACCCCACTCTGTAAATAGATTTACCGCATTTACGGCTGCAATCTGTAA 2340
QY 2341 GTGGGATGGTCTCCTAATGGAGGAGTGTTCATTTGATATAAGTTATTCACTGAGTAT 2400
DB |||||
2341 GTGGGATGGTCTCCTAATGGAGGAGTGTTCATTTGATATAAGTTATTCACTGAGTAT 2400
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGTGGCAGCA 2443
DB |||||
2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGTGGCAGCA 2443

RESULT 13

US-10-224-289-3
; Sequence 3, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; TITLE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-3

Query Match 93.7%; Score 2302; DB 7; Length 2302;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 155 CCAAGGTCTCCCCAGCACTGAGGAGCTCGCTGCTGCCCTCTTGCGCGGGAAGCAGC 214
DB |||||
1 CCAAGGTCTCCCCAGCACTGAGGAGCTCGCTGCTGCCCTCTTGCGCGGGAAGCAGC 60
QY 215 ACCAAGTTACGGCCAAAGCCCTTGCACTAGGGTCCAGAAATGCTTACACAGTCCCTGAT 274
DB |||||
61 ACCAAGTTACGGCCAAAGCCCTTGCACTAGGGTCCAGAAATGCTTACACAGTCCCTGAT 120
QY 275 GGTTCGCGCAATGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTGG 334
DB |||||
121 GGTTCGCGCAATGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTGG 180
QY 335 GGCATCGTCTAGAAAACGGTGGCCACAGCCGGGTTGTGAACCTCGGTGGCCCTTCATGCTC 394
DB |||||
181 GGCATCGTCTAGAAAACGGTGGCCACAGCCGGGTTGTGAACCTCGGTGGCCCTTCATGCTC 240
QY 395 ACTCTCCGATCCTCGTCTGCAAGGTGCAAGGACTCCACAGCGGAAATATGCTGCTACT 454
DB |||||
241 ACTCTCCGATCCTCGTCTGCAAGGTGCAAGGACTCCACAGCGGAAATATGCTGCTACT 300
QY 455 CAGTTTCTCTTCTCCTCGGTGTTGGGTCATCTTTGGCTCACCTTGCCTTTCATCATC 514
DB |||||
301 CAGTTTCTCTTCTCCTCGGTGTTGGGTCATCTTTGGCTCACCTTGCCTTTCATCATC 360
QY 515 GGCATCGGAGGAGCACAGGGCCACAGCTTCTTCTCTTTGGGATCCTCTTTTCCATC 574
DB |||||
361 GGCATCGGAGGAGCACAGGGCCACAGCTTCTTCTCTTTGGGATCCTCTTTTCCATC 420
QY 575 TGCTTCTCTGCTGCTGCTCATGCTGCTGAGCAAGCTGTCGGGGGAGGAAG 634
DB |||||
421 TGCTTCTCTGCTGCTGCTCATGCTGCTGAGCAAGCTGTCGGGGGAGGAAG 480
QY 635 CCCCCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694
DB |||||
481 CCCCCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 695 ATCGCTATTGAATATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754
DB |||||
541 ATCGCTATTGAATATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 755 CTTTTCGCTCTCGTTCGCAAGGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
DB |||||
601 CTTTTCGCTCTCGTTCGCAAGGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 815 ATGGCGCTGACCTTCTCATGCTCTCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874
DB |||||
661 ATGGCGCTGACCTTCTCATGCTCTCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 875 AGACATGGGGCCACATCTTACCTACGATGCTCTCTCCATTTGCCATCTGGGTGGCTGG 934
DB |||||
721 AGACATGGGGCCACATCTTACCTACGATGCTCTCTCCATTTGCCATCTGGGTGGCTGG 780
QY 935 ATCACCTGCTCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994
DB |||||
781 ATCACCTGCTCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 995 GCCTTGGCTGCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054
DB |||||
841 GCCTTGGCTGCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1055 CTCACAAAGCAACGAAACCCCATGGAATATCTGTTGAGGATGCTTTTGTAAACCTCAA 1114
DB |||||
901 CTCACAAAGCAACGAAACCCCATGGAATATCTCTGTTGAGGATGCTTTTGTAAACCTCAA 960
QY 1115 CTGCTGAAGAGAGCTATGGTGTGAGAGCAGAGCTTCTCTCAAGAGAAATCACTCAA 1174
DB |||||
961 CTGCTGAAGAGAGCTATGGTGTGAGAGCAGAGCTTCTCTCAAGAGAAATCACTCAA 1020
QY 1175 GGTTTTGAAGAGAGAGGAGCAGCTCTATGCCCCCTTATCCACACATTTTTCAGCTGCAG 1234
DB |||||
1021 GGTTTTGAAGAGAGAGGAGCAGCTCTATGCCCCCTTATCCACACATTTTTCAGCTGCAG 1080

Qy	1235	AACCAGGCTCCCCAAAAAGGAATTTCTCCATCCACGGGCCACAGCTTGGCCGAGCCCTTAC	1294
Db	1081	AACCAGGCTCCCCAAAAAGGAATTTCTCCATCCACGGGCCACAGCTTGGCCGAGCCCTTAC	1140
Qy	1295	AAAGCACTATGAAGTAAAGAAAGAGGGGAGCTAACTCTGCTCTGAAGAGTGGGCAAAATGC	1354
Db	1141	AAAGCACTATGAAGTAAAGAAAGAGGGGAGCTAACTCTGCTCTGAAGAGTGGGCAAAATGC	1200
Qy	1355	AGCCGGGCGGCAGACTAGCGGGGAGCTCAAGGGGATGTGGGCGGAAATCTTGAGTCTTCTG	1414
Db	1201	AGCCGGGCGGCAGACTAGCGGGGAGCTCAAGGGGATGTGGGCGGAAATCTTGAGTCTTCTG	1260
Qy	1415	AGAAAACTGTACAAGACACTACGGGAACAGTTTTGCCCTCCCTCCAGGCTCAACACAAATT	1474
Db	1261	AGAAAACTGTACAAGACACTACGGGAACAGTTTTGCCCTCCCTCCAGGCTCAACACAAATT	1320
Qy	1475	CTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTTAGAGGGCGCTGTAGTAT	1534
Db	1321	CTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTTAGAGGGCGCTGTAGTAT	1380
Qy	1535	TTTTTTTTTTTTGTCTCATCCCTTTGGATACTCTTTTAAAGTGGAGTCTCAGGCAACTCA	1594
Db	1381	TTTTTTTTTTTTGTCTCATCCCTTTGGATACTCTTTTAAAGTGGAGTCTCAGGCAACTCA	1440
Qy	1595	AGTTAGACCTTACTCTTTTGTGTTTTTTTGAACAGGATCTTGCTCTGTCCACCAG	1654
Db	1441	AGTTAGACCTTACTCTTTTGTGTTTTTTTGAACAGGATCTTGCTCTGTCCACCAG	1500
Qy	1655	GCTTGAGTGCAGTGGTGCGATCACAGCCAGTGCAGGCTCGACACCTGTGCTCAAGCAA	1714
Db	1501	GCTTGAGTGCAGTGGTGCGATCACAGCCAGTGCAGGCTCGACACCTGTGCTCAAGCAA	1560
Qy	1715	TCCTCCCATCTCATCTCCAAAGTGTGGGATGACAGGGGTGAGCCACAGTCCCGAGCC	1774
Db	1561	TCCTCCCATCTCCATCTCCAAAGTGTGGGATGACAGGGGTGAGCCACAGTCCCGAGCC	1620
Qy	1775	TAGGCCCTTAATCTTGCTGTATTTTTCCATGGACTAAGGCTCTGGTCACTCTGAGCTCAAG	1834
Db	1621	TAGGCCCTTAATCTTGCTGTATTTTTCCATGGACTAAGGCTCTGGTCACTCTGAGCTCAAG	1680
Qy	1835	CTGGCTCACACAGCTCTAGGGGCTGTCTCTTAACCTCACAGTGGGTTTTGTGAGGCTCT	1894
Db	1681	CTGGCTCACACAGCTCTAGGGGCTGTCTCTTAACCTCACAGTGGGTTTTGTGAGGCTCT	1740
Qy	1895	GTGGCCGAGAGCAGCTGCAATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACT	1954
Db	1741	GTGGCCGAGAGCAGCTGCAATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACT	1800
Qy	1955	GGCCGTGAATCTACACTGGAGCCAACTTGCTGGCACCCCGCTCCCAACCCCTCTTGCC	2014
Db	1801	GGCCGTGAATCTACACTGGAGCCAACTTGCTGGCACCCCGCTCCCAACCCCTCTTGCC	1860
Qy	2015	TGGGTAGGAGGCTAAAGATCACCCCTAAATTTTACTCATCTCTCTAGTGTGCTCCACAT	2074
Db	1861	TGGGTAGGAGGCTAAAGATCACCCCTAAATTTTACTCATCTCTCTAGTGTGCTCCACAT	1920
Qy	2075	TGGGCCTCAGAGCTCCCGACGACCAATTCACAGGTGACCCCTCTCTCTTGACCTGTCC	2134
Db	1921	TGGGCCTCAGAGCTCCCGACGACCAATTCACAGGTGACCCCTCTCTCTTGACCTGTCC	1980
Qy	2135	CCAAACTTGCTCAATTCGAGATCTAAATCTCCCCCTAGCTCTGCCAGGAAATCTTTTC	2194
Db	1981	CCAAACTTGCTCAATTCGAGATCTAAATCTCCCCCTAGCTCTGCCAGGAAATCTTTTC	2040
Qy	2195	AGACCTCACTAGCAACAGCCCGTGTCTCTGTGTCAGGAGAAATTTGTAGATCAATCTTCAC	2254
Db	2041	AGACCTCACTAGCAACAGCCCGTGTCTCTGTGTCAGGAGAAATTTGTAGATCAATCTTCAC	2100
Qy	2255	TTCAAAATTCCTGGGGCTGATCTTCTCATCTGTGACACCCCAACCTCTGTAAATAGATTT	2314
Db	2101	TTCAAAATTCCTGGGGCTGATCTTCTCATCTTGACCCCAACCTCTGTAAATAGATTT	2160

Qy	2315	ACGCATTATACGGCTGCATTTCTGTAAAGTGGGCATGGTCTCCTTAATGAGAGAGTGTTCATT	2314
Db	2161	ACGCATTATACGGCTGCATTTCTGTAAAGTGGGCATGGTCTCCTTAATGAGAGAGTGTTCATT	2220
Qy	2375	GTATATAAGCTTATTCACCTGAGTGTGCAATAAAGATGTGGTGGCCACTCTTTTCATGGTG	2434
Db	2221	GTATAATAAGTTATTTACCTTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTTCATGGTG	2280
Qy	2435	GTGGCAGCAAAAAAAAAAAAAA 2456	
Db	2281	GTGGCAGCAAAAAAAAAAAAAA 2302	
RESULT 14			
US-10-240-425-405			
; Sequence 405, Application US/10240425			
; Publication No. US20040033502A1			
; GENERAL INFORMATION:			
; APPLICANT: Williams, Amanda			
; APPLICANT: Bolland, Joseph F.			
; APPLICANT: Lord, Reginald V.			
; APPLICANT: Alvarez, Chris			
; APPLICANT: Wetzel, Jon C.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Vockley, Joseph G.			
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue			
; FILE REFERENCE: 44921-5026			
; CURRENT APPLICATION NUMBER: US/10/240,425			
; PRIOR APPLICATION NUMBER: PCT/US01/09847			
; PRIOR FILING DATE: 2002-09-30			
; PRIOR FILING DATE: 2001-03-28			
; PRIOR APPLICATION NUMBER: US 60/193,446			
; PRIOR FILING DATE: 2000-03-31			
; NUMBER OF SEQ ID NOS: 1588			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 405			
; LENGTH: 2302			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF095448			
US-10-240-425-405			
Query Match 93.7%; Score 2302; DB 8; Length 2302;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	155	CCAAGGTCCTCCCGACGACTGAGAGGCTGCCTCTGCTGCTCCCTCTTTGCGCGGGGAAGCAGC	214
Db	1	CCAAGGTCCTCCCGACGACTGAGAGGCTGCCTCTGCTGCTCCCTCTTTGCGCGGGGAAGCAGC	60
Qy	215	ACCAAGTTACGGCCAAACGCTTTGGCACTAGGGTCCAGAATGGCTACAAAGTCCCTTGAT	274
Db	61	ACCAAGTTACGGCCAAACGCTTTGGCACTAGGGTCCAGAATGGCTACAAAGTCCCTTGAT	120
Qy	275	GGTTGCGCGCAATGGCCTGAAATCCAAGTACTACAGACTTTTGTGATAAGGCTGAAGCTTGG	334
Db	121	GGTTGCGCGCAATGGCCTGAAATCCAAGTACTACAGACTTTTGTGATAAGGCTGAAGCTTGG	180
Qy	335	GGCATCGTCTTAGAAGCGTGGCCACAGCGGGTTGTGACCTCGTGGCCTTCATGCTC	394
Db	181	GGCATCGTCTTAGAAGCGTGGCCACAGCGGGTTGTGACCTCGTGGCCTTCATGCTC	240
Qy	395	ACTCTCCGATCTCGTCTCGAAGGTCGAGGACTCCAAACAGGCGAAAAATGCTGCCTACT	454
Db	241	ACTCTCCGATCTCGTCTCGAAGGTCGAGGACTCCAAACAGGCGAAAAATGCTGCCTACT	300
Qy	455	CAGTTTCTCTTCTCTCTGGGTGTGTTGGGCATCTTTGGCCTCACCTTCGCTTCATCATC	514
Db	301	CAGTTTCTCTTCTCTCTGGGTGTGTTGGGCATCTTTGGCCTCACCTTCGCTTCATCATC	360
Qy	515	GGACTGGAGCGGAGCAGAGGCCCCACACGCTTCTTCTCTTTGGGATCCCTCTTTTTCATC	574

[illegible]

RESIT.T 15

US-10-775-920-9

US-10-773-920-9
: Sequence 9. Application US/10775920

Publication No. US20040175744A1

; PUBLICATION NO: US2004
: GENERAL INFORMATION:

APPLICANT: Mergen Ltd

APPLICANT: Mergem Ltd
TITLE OF INVENTION: Po

NOT INVENTION OF LITTLE

1. TITLE OF INVENTION: _____

FILE PREFERENCE: Merges

```

; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: IIS/10/775 920

```

; CURRENT APPLICATION NUMBER: US/1
 : CURRENT FILING DATE: 2004-03-10

; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,800

PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2003-03-13

; PRIOR FILING DATE: 2003-0

; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 9									
; LENGTH: 2302									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-775-920-9									
Query Match 93.7%; Score 2302; DB 9; Length 2302;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	155	CCAAGTCTCC	CCCAAGCACTG	AGGAGCTCG	CCCTGCTG	CCCTCTTTG	CGCGGGAAG	CAGC	214
DB	1	CCAAGTCTCC	CCCAAGCACTG	AGGAGCTCG	CCCTGCTG	CCCTCTTTG	CGCGGGAAG	CAGC	60
QY	215	ACCAAGTTC	CACGGCCAA	CGCTTGG	CACTAGG	GTCCAGAT	GGCTACA	CAGTCCCTGAT	274
DB	61	ACCAAGTTC	CACGGCCAA	CGCTTGG	CACTAGG	GTCCAGAT	GGCTACA	CAGTCCCTGAT	120
QY	275	GGTTGCGCA	ATGGCCTG	AAATCCAA	GTACTAC	AGACTTTT	GTGATA	AAAGCTGAAGCTTGG	334
DB	121	GGTTGCGCA	ATGGCCTG	AAATCCAA	GTACTAC	AGACTTTT	GTGATA	AAAGCTGAAGCTTGG	180
QY	335	GGCATCGTC	CTAGAAAC	CGGTGGC	CAAGCCGGG	GTGTGAC	CTCGGTGGC	CTTCATGCTC	394
DB	181	GGCATCGTC	CTAGAAAC	CGGTGGC	CAAGCCGGG	GTGTGAC	CTCGGTGGC	CTTCATGCTC	240
QY	395	ACTCTCCGA	TCTCGTCTG	CAAGGTG	CAAGGACT	CCAA	CAGGGGAAA	AAATGCTGCCTACT	454
DB	241	ACTCTCCGA	TCTCGTCTG	CAAGGTG	CAAGGACT	CCAA	CAGGGGAAA	AAATGCTGCCTACT	300
QY	455	CAGTTTCTC	TCTCCTCG	TGGGTGT	TGGGCAT	CTTTGGG	CTCACC	TTTCGCCCTTCATCATC	514
DB	301	CAGTTTCTC	TCTCCTCG	TGGGTGT	TGGGCAT	CTTTGGG	CTCACC	TTTCGCCCTTCATCATC	360
QY	515	GGAGTGGAC	GGGAGCAC	AGGGGCC	ACACGCTT	CTCTTTG	GGGATC	CTCTTTTCCATC	574
DB	361	GGAGTGGAC	GGGAGCAC	AGGGGCC	ACACGCTT	CTCTTTG	GGGATC	CTCTTTTCCATC	420
QY	575	TGCTTCTCT	CGCTGCTAT	GTCTAGT	CTGACCA	AGCTGCT	CGGGGGAGG	AAAG	634
DB	421	TGCTTCTCT	CGCTGCTAT	GTCTAGT	CTGACCA	AGCTGCT	CGGGGGAGG	AAAG	480
QY	635	CCCTTTCC	TCTGTGTG	ATCTGGG	CTGCGG	CTTCA	GGCTAGT	CCAGGATGTT	694
DB	481	CCCTTTCC	TCTGTGTG	ATCTGGG	CTGCGG	CTTCA	GGCTAGT	CCAGGATGTT	540
QY	695	ATCGCTATT	GGAATAT	TTGTCTG	ACCATGA	ATAGGAC	CAACG	TCATATGCTCTTTTCTGAG	754
DB	541	ATCGCTATT	GGAATAT	TTGTCTG	ACCATGA	ATAGGAC	CAACG	TCATATGCTCTTTTCTGAG	600
QY	755	CTTTCCGCT	CTCTGCG	CAATGA	AGACTTT	GTCTCTG	CTCA	CCCTACCTCTCTCTCTTG	814
DB	601	CTTTCCGCT	CTCTGCG	CAATGA	AGACTTT	GTCTCTG	CTCA	CCCTACCTCTCTCTCTTG	660
QY	815	ATGGGCTG	ACCTTCT	CTATG	CTCTT	CACTT	CTGTG	GTTCCTTACGGGCTGGAAG	874
DB	661	ATGGGCTG	ACCTTCT	CTATG	CTCTT	CACTT	CTGTG	GTTCCTTACGGGCTGGAAG	720
QY	875	AGACATGG	GGGCCCA	CACTAC	CTACG	ATGCT	CTCTT	CCATTGGCCATCTGGGTGGCCTGG	934
DB	721	AGACATGG	GGGCCCA	CACTAC	CTACG	ATGCT	CTCTT	CCATTGGCCATCTGGGTGGCCTGG	780
QY	935	ATCACCTGT	CTATGCT	CTTGA	CTTTG	ACCGAG	GTGGAT	GACACCATCTCCTCAGCTCC	994
DB	781	ATCACCTGT	CTATGCT	CTTGA	CTTTG	ACCGAG	GTGGAT	GACACCATCTCCTCAGCTCC	840
QY	995	GCCTTGGC	TGCCAA	TGGGTG	TTCTG	TGGCTT	ATGTT	AGTCCCGAGTTTGGCTG	1054
DB	841	GCCTTGGC	TGCCAA	TGGGTG	TTCTG	TGGCTT	ATGTT	AGTCCCGAGTTTGGCTG	900
QY	1055	CTCACAAAG	CACGAA	CCCACTG	GAATAT	CTCTG	TGTGAGG	ATGCTTTCTGTAAACCTCAA	1114

901	CTCACAAAGCAAAGAAACCCCAT	TGGATTATCTCTGTTGAGGATGCTTTCTGTAAACCTCAA	960
1115	CTCGTGAAGAAGAGCTATG	TGTGTGGAGAACAGAGCCCTACTCTCAAGAGGAAATCACTCAA	1174
961	CTCGTGAAGAAGAGCTATG	TGTGTGGAGAAACAGAGCCCTACTCTCAAGAGGAAATCACTCAA	1020
1175	GGTTTTGAAGAGACAGGGGAC	ACGCTCTATGCCCCCTATTTCCACACATTTTCAGCTGCAG	1234
1021	GGTTTTGAAGAGACAGGGGAC	ACGCTCTATGCCCCCTATTTCCACACATTTTCAGCTGCAG	1080
1235	AACAGCGCTCCC	CAAAAGGAATTTCTCATCCACGGGCCCAACGCTTGGCGGAGCCCTTAC	1294
1081	AACAGCGCTCCC	CAAAAGGAATTTCTCCATCCACGGGCCCAACGCTTGGCGGAGCCCTTAC	1140
1295	AAAGACTATGAAGTAAAGAA	GAGGCGAGCTAACTCTGTCTCTGAAGAGTGGGACAAATGC	1354
1141	AAAGACTATGAAGTAAAGAA	GAGGCGAGCTAACTCTGTCTCTGAAGAGTGGGACAAATGC	1200
1355	AGCGGGCGGCAGATCTAG	CGGGAGCTCAAAAGGATGTGGGGGAAATCTTTGAGTCTTCTG	1414
1201	AGCGGGCGGCAGATCTAG	CGGGAGCTCAAAAGGATGTGGGGGAAATCTTTGAGTCTTCTG	1260
1415	AGAAACTGTAC	AAGACACTACCGGAAACAGTTTGCCTCCCTCCAGCGCTCAACCCACAAT	1474
1261	AGAAACTGTAC	AAGACACTACCGGAAACAGTTTGCCTCCCTCCAGCGCTCAACCCACAAT	1320
1475	CTTCCATGCTGGGGCTGAT	GTGGGCTAGTAAAGACTCCAGTCTTTAGAGGGGCTGTAGTAT	1534
1321	CTTCCATGCTGGGGCTGAT	GTGGGCTAGTAAAGACTCCAGTCTTTAGAGGGGCTGTAGTAT	1380
1535	TTTTTTTTTTTTTGTCTCA	TCTTTTGGATACCTCTTTTAAAGTGGGAGTCTCAGGCAACTCA	1594
1381	TTTTTTTTTTTTTGTCTCA	TCTTTTGGATACCTCTTTTAAAGTGGGAGTCTCAGGCAACTCA	1440
1595	AGTTTAGACCTTA	CTCTTTTGTGTTTTTAAACAGGATCTTGCTCTGTCAACCAG	1654
1441	AGTTTAGACCTTA	CTCTTTTGTGTTTTTAAACAGGATCTTGCTCTGTCAACCAG	1500
1655	GCCTGAGTGCAGTGGTG	CGATCACAGCCAGTGCAGCCTCGACCACTGTGCTCAAGCAA	1714
1501	GCCTGAGTGCAGTGGTG	CGATCACAGCCAGTGCAGCCTCGACCACTGTGCTCAAGCAA	1560
1715	TCCTCCCATCTCOATCT	CTCCAAAGTGTGGGATGACAGGGGTGAGCCACAGCTCCAGCC	1774
1561	TCCTCCCATCTCCATCT	CTCCAAAGTGTGGGATGACAGGGGTGAGCCACAGCTCCAGCC	1620
1775	TAGGCCCTTAATCTGCT	GTATTTTCCATGAGCTAAAGGTCTGGTCTCATCTGAGCTCACG	1834
1621	TAGGCCCTTAATCTGCT	GTATTTTCCATGAGCTAAAGGTCTGGTCTCATCTGAGCTCACG	1680
1835	CTGGCTCACACAGCTCT	TAGGGGCTGTCTCTTAACCTCACAGTGGGTTTTGTGAGGCTCT	1894
1681	CTGGCTCACACAGCTCT	TAGGGGCTGTCTCTCTTAACCTCACAGTGGGTTTTGTGAGGCTCT	1740
1895	GTGGCCCAAGCAGACG	CTGCTATCTGAGCAAAATAAGCAAAAGCCCTCTCTCAGGCCACT	1954
1741	GTGGCCCAAGCAGACG	CTGCTATCTGAGCAAAATAAGCAAAAGCCCTCTCTCAGGCCACT	1800
1955	GGCCTGAAATCTAC	ACTGGAAGCCAACTTGTGGCACCCCGCTCCCAACCCCTCTTGCC	2014
1801	GGCCTGAAATCTAC	ACTGGAAGCCAACTTGTGGCACCCCGCTCCCAACCCCTCTTGCC	1860
2015	TGGGTAGGAGGCTAA	AGATCAACCTTAAATTTACTCATCTCTAGTGTGCCTCACAT	2074
1861	TGGGTAGGAGGCTAA	AGATCAACCTTAAATTTACTCATCTCTAGTGTGCCTCACAT	1920
2075	TGGGCCCTCAGCAGCT	CCCCAGCACCAATTCACAGGTCAACCCCTCTCTTGTGCACTGTCC	2134
1921	TGGGCCCTCAGCAGCT	CCCCAGCACCAATTCACAGGTCAACCCCTCTCTTGTGCACTGTCC	1980
2135	CCAAACTTGTGTCAAT	TCGAGATCTAATCTCCCGCTACGCTCTGCGAGGAATCTTTTC	2194
1981	CCAAACTTGTGTCAAT	TCGAGATCTAATCTCCCGCTACGCTCTGCGAGGAATCTTTTC	2040

Qy	2195	AGACCTCACTAGCAACAAGCCCGTTGCTCCTTGTGTCAGGAGAAATTTGTAGATCAATTCTCAC	2254
Db	2041	AGACCTCACTAGCAACAAGCCCGTTGCTCCTTGTGTCAGGAGAAATTTGTAGATCAATTCTCAC	2100
Qy	2255	TTCAAATTCCTGGGCTGATCTCTCTCATCTTGCACCCCAACCTCTGTAATAGATT	2314
Db	2101	TTCAAATTCCTGGGCTGATCTCTCTCATCTTGCACCCCAACCTCTGTAATAGATT	2160
Qy	2315	ACCGCATTTACGGCTGCAATTTCTGTAAGTGGGCATGGTCTCTAAATGGAGAGTGTTCATT	2374
Db	2161	ACCGCATTTACGGCTGCAATTTCTGTAAGTGGGCATGGTCTCTAAATGGAGAGTGTTCATT	2220
Qy	2375	GTATAATAAGTTATTACCTGAGTATGCAATAAAGATGTGGCCACTCTTTTCATGGTG	2434
Db	2221	GTATAATAAGTTATTACCTGAGTATGCAATAAAGATGTGGCCACTCTTTTCATGGTG	2280
Qy	2435	GTGGCAGCAAAAAAAAAAAAAA	2456
Db	2281	GTGGCAGCAAAAAAAAAAAAAA	2302

Search completed: June 10, 2006, 15:59:14
Job time : 3009 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 14:55:02 ; Search time 1403 Seconds

(without alignments)
12205.162 Million cell updates/sec

Title: US-10-600-816-2

Perfect score: 2456

Sequence: 1 ataacagcatgaagtgcggt.....ggcagcaaaaaaaaaaaaaa 2456

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_8:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005as:*
- 15: Geneseqn2005bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2456	100.0	2456	8	AB242832 Human G p
2	2456	100.0	2456	11	ADN39301 Cancer/an
3	2456	100.0	2456	12	AD128459 Human GPC
4	2456	100.0	2456	13	ADR48221 Human ret
5	2456	100.0	2456	13	ACN39388 Tumour-as
6	2456	100.0	2456	13	ADU06126 Novel bro
7	2456	100.0	2456	14	AEA00088 Human TAT
8	2456	100.0	2456	14	AEA00508 Human TAT
9	2456	100.0	2456	14	AED47480 Retinoic
10	2448	99.7	2456	12	AD128525 Human GPC
11	2443	99.5	4239	11	ACN89274 Breast ca
12	2443	99.5	6730	13	ADX97494 Pancreati
13	2441.6	99.4	2456	10	ACC72695 Human can
14	2439.8	99.3	2456	4	AHH14688 Human
15	2437	99.2	2456	13	ADR43392 Human bre
16	2302	93.7	2302	6	ABT10173 Human bre
17	2302	93.7	2302	8	ACC58386 Human GPC
17	2302	93.7	2302	10	ADD93240 RAIGI cod

19	2302	93.7	2302	10	ADL24773	Adl24773 Intestina
20	2288.8	93.2	3371	4	AAH72766	Aah72766 Human cer
21	2286	93.1	2316	10	ABT31923	Abt31923 Human bre
22	2260.4	92.0	2593	6	ABQ54954	Abq54954 Human ova
23	1701.2	69.3	1718	3	AAZ90056	Aaz90056 Hydrophob
24	1601.8	65.2	1619	4	AAF58615	Aaf58615 Human REC
25	1441.4	58.7	1460	14	AEA00112	Aea00112 Human TAT
26	1441.4	58.7	1460	14	AEA00632	Aea00632 Human TAT
27	1177.6	47.9	2347	14	ADZ88660	Adz88660 Breast sp
28	1126	45.8	1212	3	AAZ61776	Aaz61776 cDNA enco
29	1126	45.8	1212	4	AAC99709	Aac99709 Skin cell
30	1126	45.8	1212	6	ABL34861	Ab134861 Human CDN
31	1082.6	44.1	1114	2	AAZ97687	Aaz97687 Extended
32	1082.6	44.1	1114	12	ADP18954	Adp18954 Human sec
33	1074	43.7	1074	12	ADQ30035	Adq30035 Human GPC
34	1067.8	43.5	1788	10	ADF70573	Adf70573 Orphan re
35	1031.6	42.0	1228	8	ACA04775	AcA04775 cDNA enco
36	932.6	38.0	1000	4	AAI21778	Aai21778 Probe #11
37	932.6	38.0	1000	4	ABA66850	AbA66850 Human foe
38	932.6	38.0	1000	4	AAI47061	Aai47061 Probe #15
39	932.6	38.0	1000	4	ABA48935	AbA48935 Human bre
40	932.6	38.0	1000	4	ABA33918	AbA33918 Probe #12
41	932.6	38.0	1000	4	AAK41008	AaK41008 Human bon
42	932.6	38.0	1000	4	AAK15285	AaK15285 Human bra
43	932.6	38.0	1000	4	ABS40600	AbS40600 Human liv
44	932.6	38.0	1000	5	AAI07463	Aai07463 Probe #74
45	932.6	38.0	1000	6	ABS14980	AbS14980 Human gen

ALIGNMENTS

RESULT 1

AB242832

ID AB242832 standard; DNA; 2456 BP.

XX

AC AB242832;

XX

DT 04-MAR-2003 (first entry)

XX

DE Human G protein-coupled receptor RAIGI nucleotide SEQ ID NO:453.

XX

KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;

KW G protein-coupled receptor modulator; antibody; immune-related disease;

KW growth-related disease; cell regeneration-related disease; AIDS; cancer;

KW immunological-related cell proliferative disease; autoimmune disease;

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;

KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;

KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;

KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;

KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;

KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

KW ulcer; gene; ds.

XX

XX Homo sapiens.

XX

PN WO200261087-A2.

XX

PD 08-AUG-2002.

XX

PF 19-DEC-2001; 2001WO-US050107.

XX

PR 19-DEC-2000; 2000US-0257144P.

XX

PA (LIFE-) LIFESPAN BIOSCIENCES INC.

XX

PI Burmer GC, Roush CL, Brown JF;

XX

DR WPI; 2003-046718/04.

DR

DR P-PSDB; ABP81984.

XX

PT New isolated antigenic peptides e.g., for G protein-coupled receptors

PT (GPCR), useful for diagnosing and designing drugs for treating conditions

in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP2019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related diseases, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention

Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 2456; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PT	1	ATACAGCATGAAGTGGCGGTGAACTGGAAATAGGCGTGTCTCTCCCTGCACCCCTCCCCC	60
PT	1	ATACAGCATGAAGTGGCGGTGAACTGGAAATAGGCGTGTCTCTCCCTGCACCCCTCCCCC	60
XX	61	TCCTTGTCCCTGTCTACCCCTCGCTGTTCCCTCCCTCCGCGAGGCGCGCTTTATA	120
XX	61	TCCTTGTCCCTGTCTACCCCTCGCTGTTCCCTCCCTCCGCGAGGCGCGCTTTATA	120
QY	121	ACAACTGCTCAGAGTGGCGGCGGATAGTGTCCAAAGGTCTCCCGCCAGCACTCAGGAG	180
DB	121	ACAACTGCTCAGAGTGGCGGCGGATAGTGTCCAAAGGTCTCCCGCCAGCACTCAGGAG	180
QY	181	CTCGCTGCTGCTCTTTGCGCGGGGAGAGCAAGTTCAGGTTTCAGCGGCAAGCTTGGC	240
DB	181	CTCGCTGCTGCTCTTTGCGCGGGGAGAGCAAGTTCAGGTTTCAGCGGCAAGCTTGGC	240
QY	241	ACTAGGGTCCAGAAAGGCTTACAAAGTCCCTGATGTTGGTGGCGCAATGGCTGAAATCCAA	300
DB	241	ACTAGGGTCCAGAAAGGCTTACAAAGTCCCTGATGTTGGTGGCGCAATGGCTGAAATCCAA	300
QY	301	GTACTACAGACTTTGTGATAAGGCTTGAAGCTTGGGGCATCGTCTAGAAACCGTGGCCAC	360
DB	301	GTACTACAGACTTTGTGATAAGGCTTGAAGCTTGGGGCATCGTCTAGAAACCGTGGCCAC	360
QY	361	AGCGGGGTGTGACCTCGTGGGCTTCAATGCTCACTCTCCCGATCTCTCCGATCTCTGGG	420
DB	361	AGCGGGGTGTGACCTCGTGGGCTTCAATGCTCACTCTCCCGATCTCTCCGATCTCTGGG	420
QY	421	GCAGACTCCAAAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTCTCTCTCTCTCTCTCT	480
DB	421	GCAGACTCCAAAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTCTCTCTCTCTCTCTCT	480
QY	481	GGGCACTTTGGGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT	540
DB	481	GGGCACTTTGGGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT	540
QY	541	AGCTTCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTCTCTCTCTCTCTCTCTCT	600
DB	541	AGCTTCTTCTCTCTTGGGATCCTCTTTTCCATCTGCTTCTCTCTCTCTCTCTCTCTCTCT	600
QY	601	TCTAGCTGTGACCAAGCTCGTCCGGGGAGGAGCCCTTTTCCCTTGTGTGATTTCTGGG	660
DB	601	TCTAGCTGTGACCAAGCTCGTCCGGGGAGGAGCCCTTTTCCCTTGTGTGATTTCTGGG	660
QY	661	TCTGGCCGTGGGCTTTCAGCTAGTCCAGGATGTTATCGTATTGAATATTGTCTCTGAC	720
DB	661	TCTGGCCGTGGGCTTTCAGCTAGTCCAGGATGTTATCGTATTGAATATTGTCTCTGAC	720
QY	721	CATGAATAGGACCAAGCTCAATGTCTTTCTGAGCTTTCCGCTCTCTGCGCAATGAAGA	780
DB	721	CATGAATAGGACCAAGCTCAATGTCTTTCTGAGCTTTCCGCTCTCTGCGCAATGAAGA	780
QY	781	CTTCTGCT	840
DB	781	CTTCTGCT	840
QY	841	CTTCACT	900
DB	841	CTTCACT	900
QY	901	GATGCT	960
DB	901	GATGCT	960
QY	961	TGACCGCAGGTGGGATGACACCATCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020
DB	961	TGACCGCAGGTGGGATGACACCATCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020
QY	1021	CCTGTTGGCTTATGTTAGTCCGAGTTTGGCTGTCTCAAAAGCAACGAAACCCCATGA	1080
DB	1021	CCTGTTGGCTTATGTTAGTCCGAGTTTGGCTGTCTCAAAAGCAACGAAACCCCATGA	1080
QY	1081	TTATCTCTGTTGAGGATGCTTCTGTAACCTCAACTGCTGAGAGAGCTATGTTGTGA	1140
DB	1081	TTATCTCTGTTGAGGATGCTTCTGTAACCTCAACTGCTGAGAGAGCTATGTTGTGA	1140
QY	1141	GAAACAGGCTTACTCTCAAGAGGAAATCACTCAAGTTTGAAGAGACAGGGAACACGCT	1200
DB	1141	GAAACAGGCTTACTCTCAAGAGGAAATCACTCAAGTTTGAAGAGACAGGGAACACGCT	1200
QY	1201	CTATGCCCCCTATTCCACACATTTTTCAGTGTGAGAAACAGGCTTCCCAAAAGAAATCTC	1260
DB	1201	CTATGCCCCCTATTCCACACATTTTTCAGTGTGAGAAACAGGCTTCCCAAAAGAAATCTC	1260
QY	1261	CATCCACGCGGCGGCGGCTTGGCGAGCCCTTACAAAGCACTATGAAGTAAAGAAAGAGG	1320
DB	1261	CATCCACGCGGCGGCGGCTTGGCGAGCCCTTACAAAGCACTATGAAGTAAAGAAAGAGG	1320
QY	1321	CAGCTTAACTCTGTCTGTAAGAGTGGGCAAAATGCAAGCGGGCGGCGAGATCTAGCGGAGC	1380
DB	1321	CAGCTTAACTCTGTCTGTAAGAGTGGGCAAAATGCAAGCGGGCGGCGAGATCTAGCGGAGC	1380
QY	1381	TCAAAGGATGTGGGCGGAAATCTTGAAGTCTTCTGAGAAAACTGTACAGACACACTACGGGA	1440
DB	1381	TCAAAGGATGTGGGCGGAAATCTTGAAGTCTTCTGAGAAAACTGTACAGACACACTACGGGA	1440
QY	1441	ACAGTTTCCCTCCCTCCAGCTCAACCAATCTTCCATGCTGGGCTGATGTGGCT	1500
DB	1441	ACAGTTTCCCTCCCTCCAGCTCAACCAATCTTCCATGCTGGGCTGATGTGGCT	1500
QY	1501	AGTAAGACTCCAGTTCTTAGAGGCGCTAGTATTTTTTTTTTTTTTTTTTTTGTCTCATCTTTGG	1560
DB	1501	AGTAAGACTCCAGTTCTTAGAGGCGCTAGTATTTTTTTTTTTTTTTTTTTTGTCTCATCTTTGG	1560
QY	1561	ATATCTTTTAAAGTGGGAGTCTAGGCAACTCAAGTTTAGACCTTACTCTTTTGTGTT	1620
DB	1561	ATATCTTTTAAAGTGGGAGTCTAGGCAACTCAAGTTTAGACCTTACTCTTTTGTGTT	1620
QY	1621	GTTTTTTGAACAGGATCTTGTCTGTCTACCCAGGCTTTGAGTGCAGTGGTGCATCAG	1680

Db 1621 GTTTTGAACAGGATCTTGCTCTGTCTCAGCCAGGCTTGAGTGCAGTGGTGCATCAG 1680
Qy 1681 CCAGTGCAGCTCGACCCAGCTGTGTCTCAAGCAATCTCTCCATCTCTCCAAAGTG 1740
Db 1681 CCAGTGCAGCTCGACCCAGCTGTGTCTCAAGCAATCTCTCCATCTCTCCAAAGTG 1740
Qy 1741 CTGGGATGACAGGCTGAGCCAGCTCCAGCTCCAGCTTAAATCTTGTGTTATTTT 1800
Db 1741 CTGGGATGACAGGCTGAGCCAGCTCCAGCTCCAGCTTAAATCTTGTGTTATTTT 1800
Qy 1801 CCATGGAATAAAGGTCTGTGTCATCTGAGCTCAGCTGGTTCACACAGCTCTAGGGCCCTG 1860
Db 1801 CCATGGAATAAAGGTCTGTGTCATCTGAGCTCAGCTGGTTCACACAGCTCTAGGGCCCTG 1860
Qy 1861 CTCCTCTAATCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGCTGCATATCT 1920
Db 1861 CTCCTCTAATCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGCTGCATATCT 1920
Qy 1921 GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCCTGAATCTACACTGGAAGCCAAAC 1980
Db 1921 GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCCTGAATCTACACTGGAAGCCAAAC 1980
Qy 1981 TTGCTGGCACCCCGCTCCCAACCCCTTCTTGCTGGTAGGAGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCCGCTCCCAACCCCTTCTTGCTGGTAGGAGGCTAAAGATCACCC 2040
Qy 2041 TAAATTTACTATCTCTAGTCTGCTCAGCTCAGCTGGGCTCAGAGCTCCCGAGCACCA 2100
Db 2041 TAAATTTACTATCTCTAGTCTGCTCAGCTCAGCTGGGCTCAGAGCTCCCGAGCACCA 2100
Qy 2101 ATTACAGCTCACCCCTCTCTTCTGCACTGCTCCCAAACTTCTCTCAATTCGGAGATC 2160
Db 2101 ATTACAGCTCACCCCTCTCTTCTGCACTGCTCCCAAACTTCTCTCAATTCGGAGATC 2160
Qy 2161 TAATCTCCCTTACGCTCTGCCAGGAATCTTTTTCAGACCTCTCAGTACCAAGCCCGGTTG 2220
Db 2161 TAATCTCCCTTACGCTCTGCCAGGAATCTTTTTCAGACCTCTCAGTACCAAGCCCGGTTG 2220
Qy 2221 CTCCTTGTGAGGAAATTTGTAGATCATCTCTCAATTCATTAATTAAGTTATTCACTGAGTAT 2280
Db 2221 CTCCTTGTGAGGAAATTTGTAGATCATCTCTCAATTCATTAATTAAGTTATTCACTGAGTAT 2280
Qy 2281 CTCATCTTGCACCCCACTCTCTTAATAGATTTTACCGATTTACCGCTGCAATTCGTAA 2340
Db 2281 CTCATCTTGCACCCCACTCTCTTAATAGATTTTACCGATTTACCGCTGCAATTCGTAA 2340
Qy 2341 GTGGGATGCTCTCTTAATGGAGGAGTGTTCATTTATTAATTAAGTTATTCACTGAGTAT 2400
Db 2341 GTGGGATGCTCTCTTAATGGAGGAGTGTTCATTTATTAATTAAGTTATTCACTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGGGCACTCTTTTCATGTGGTGGGAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGGCACTCTTTTCATGTGGTGGGAGCAAAAAA 2456

RESULT 2
ADN39301

ID ADN39301 standard; cDNA; 2456 BP.

AC ADN39301;

DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:619.

XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnarary; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-032464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 08-FEB-2002; 2002US-0355250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-0368809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 05-JUN-2002; 2002US-0386614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-0397775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

WPI; 2003-468649/44.
P-PSDB; ADN39302.

Determining the presence or absence of a pathological cell in a patient,
useful for diagnosing, prognosing or treating cancer, comprises detecting
a nucleic acid in a biological sample.

Claim 8; SEQ ID NO 619; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
whose expression is upregulated or downregulated in specific cancers or
other diseases such as angiogenic or fibrotic disorders, and to methods
of determining the presence or absence of a pathological cell in a
patient by detecting a nucleic acid at least 80% identical to those of
the invention or by detecting a polypeptide of the invention. The
invention also relates to expression vectors and host cells comprising a
nucleic acid of the invention; antibodies which specifically bind a
polypeptide of the invention; use of such antibodies for drug targeting;
and methods of screening for modulators of activity or expression of the
polypeptides and nucleic acids. The nucleic acids, polypeptides,
antibodies and methods are useful for diagnosing, prognosing and treating
cancer and other conditions such as psoriasis, ischaemia, heart disease,
atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
neovascularisation syndromes, scarring and uterine fibroids. They may
also be useful in wound healing and in contraception. The present
sequence represents a nucleic acid sequence of the invention.

Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 2456; DB 11; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATACAGCATGAAGTGCCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCTCCCC 60
Db 1 ATACAGCATGAAGTGCCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCTCCCC 60
Qy 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGCTCCCTCCCTCCGCGAGGCGCGCTTTATA 120
Db 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGCTCCCTCCCTCCGCGAGGCGCGCTTTATA 120

QY	121	ACAACTGCTCAGAGTCGAGGCGGGAATGCTGTCCAAAGTCTCCCCAGCACTGAGGAG	180
Db	121	ACAACTGCTCAGAGTCGAGGCGGGAATGCTGTCCAAAGTCTCCCCAGCACTGAGGAG	180
QY	181	CTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGCACCAGATTACGGCCAAAGCCTTGGC	240
Db	181	CTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGCACCAGATTACGGCCAAAGCCTTGGC	240
QY	241	ACTAGGGTCCAGAAATGGCTACAACAGTCCCTGATGTTGCCGCAATGGCTGAAATCCAA	300
Db	241	ACTAGGGTCCAGAAATGGCTACAACAGTCCCTGATGTTGCCGCAATGGCTGAAATCCAA	300
QY	301	GTACTACAGACTTTGTGATPAAGCTGAAAGCTTTGGGCACTCGTCCTAGAAGACGGTGGCCAC	360
Db	301	GTACTACAGACTTTGTGATPAAGCTGAAAGCTTTGGGCACTCGTCCTAGAAGACGGTGGCCAC	360
QY	361	AGCCGGGTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTGCTGCAAGT	420
Db	361	AGCCGGGTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTGCTGCAAGT	420
QY	421	GCAGGACTCCAAACAGCGCAAAATGCTGCTACTCAGATTCTCTCTCTCTGGGTGTGT	480
Db	421	GCAGGACTCCAAACAGCGCAAAATGCTGCTACTCAGATTCTCTCTCTCTGGGTGTGT	480
QY	481	GGGCATCTTTGGCCTCACCTTCGCCCTTCATCGGACTGGAAGGAGCAGAGGCCAC	540
Db	481	GGGCATCTTTGGCCTCACCTTCGCCCTTCATCGGACTGGAAGGAGCAGAGGCCAC	540
QY	541	AGCTCTCTCTCTTTGGGATCCTCTTTTCCATCTGCTCTCTCGCTGCTGCTCATGC	600
Db	541	AGCTCTCTCTCTTTGGGATCCTCTTTTCCATCTGCTCTCTCGCTGCTGCTCATGC	600
QY	601	TGTCAGTCTGACCAAGCTCGTCGGGGAGGAAGCCCTTTCCCTGTTGGTGATCTGGG	660
Db	601	TGTCAGTCTGACCAAGCTCGTCGGGGAGGAAGCCCTTTCCCTGTTGGTGATCTGGG	660
QY	661	TCGCGCGTGGCTCAGCTAGTCAGGATGTTATCGCTATGAAATATATTTGCTCTGAC	720
Db	661	TCGCGCGTGGCTCAGCTAGTCAGGATGTTATCGCTATGAAATATATTTGCTCTGAC	720
QY	721	CATGATAGGACCAAGTCAATGCTCTTTCTCAGCTTTCCGCTCTCCGTCGCAATGAAGA	780
Db	721	CATGATAGGACCAAGTCAATGCTCTTTCTCAGCTTTCCGCTCTCCGTCGCAATGAAGA	780
QY	781	CTTTGCTCTCTGCTCACCTACGTCCTCTTTTGATGGCGCTCAACCTTCTCATGTCCTC	840
Db	781	CTTTGCTCTCTGCTCACCTACGTCCTCTTTTGATGGCGCTCAACCTTCTCATGTCCTC	840
QY	841	CTTCACCTTCTGTGGTTCCTTCACGGCTGGGAAGACATGGGGCCCAATCTACCTCAC	900
Db	841	CTTCACCTTCTGTGGTTCCTTCACGGCTGGGAAGACATGGGGCCCAATCTACCTCAC	900
QY	901	GATGCTCTCTCCATGCTGCTGGTGGCTGGATCACTCCCTGCTCATGCTTCTGACTT	960
Db	901	GATGCTCTCTCCATGCTGCTGGTGGCTGGATCACTCCCTGCTCATGCTTCTGACTT	960
QY	961	TGACCGCAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGT	1020
Db	961	TGACCGCAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGT	1020
QY	1021	CTGTGTGGCTTATGTTAGTCCGAGTTTGGCTGTCTCAAAAGCAAGCAAGCAAGCAAGCA	1080
Db	1021	CTGTGTGGCTTATGTTAGTCCGAGTTTGGCTGTCTCAAAAGCAAGCAAGCAAGCAAGCA	1080
QY	1081	TTATCTGTTGAGGATGCTTCTGTAACCTCACTCGTGAAGAGAGCTATGTTGTTGA	1140
Db	1081	TTATCTGTTGAGGATGCTTCTGTAACCTCACTCGTGAAGAGAGCTATGTTGTTGA	1140
QY	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGAGCACGCT	1200
Db	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGAGCACGCT	1200

QY	1201	CTATGCCCTTATTTCCACACATTTTTCAGCTGCAGAACCAAGCCTCCCAAAAAGAAATTTCTC	1260
Db	1201	CTATGCCCTTATTTCCACACATTTTTCAGCTGCAGAACCAAGCCTCCCAAAAAGAAATTTCTC	1260
QY	1261	CATCCACGGGCCCCACGCTTGGCCGAGCCTTACAAAGACTATGAAGTAAAGAAAGAGGG	1320
Db	1261	CATCCACGGGCCCCACGCTTGGCCGAGCCTTACAAAGACTATGAAGTAAAGAAAGAGGG	1320
QY	1321	CAGCTAACTCTCTCCTGAAGAGTGGGACAAATGACGCGGGCGGAGATCTTAGCGGAGC	1380
Db	1321	CAGCTAACTCTCTCCTGAAGAGTGGGACAAATGACGCGGGCGGAGATCTTAGCGGAGC	1380
QY	1381	TCAAAAGGATGTGGGCGAAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA	1440
Db	1381	TCAAAAGGATGTGGGCGAAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA	1440
QY	1441	ACAGTTTGCCTCCCTCCAGCCTCAACCAATTTCTTCCATCTGGGGCTGATGTGGGT	1500
Db	1441	ACAGTTTGCCTCCCTCCAGCCTCAACCAATTTCTTCCATCTGGGGCTGATGTGGGT	1500
QY	1501	AGTAAGACTCCAAGTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTCTTTGG	1560
Db	1501	AGTAAGACTCCAAGTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTCTTTGG	1560
QY	1561	ATACTTTTAAAGTGGGAGTCTCAGGCACTCAAGTTTAGACCTTACTCTTTTGT	1620
Db	1561	ATACTTTTAAAGTGGGAGTCTCAGGCACTCAAGTTTAGACCTTACTCTTTTGT	1620
QY	1621	GTTTTTTGAACAGAGTCTTGCTGTCAACCAAGCTTGAGTGCAGTGTGGATCACAG	1680
Db	1621	GTTTTTTGAACAGAGTCTTGCTGTCAACCAAGCTTGAGTGCAGTGTGGATCACAG	1680
QY	1681	CCAGTGCAGCTCGACCACTGTCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG	1740
Db	1681	CCAGTGCAGCTCGACCACTGTCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG	1740
QY	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTGCTGTTATTTT	1800
Db	1741	CTGGGATGACAGGCTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTGCTGTTATTTT	1800
QY	1801	CCATGGACTAAAGTCTGGTCTCATCTGAGCTCAGCTCGCTCACACAGCTCTAGGGGCTG	1860
Db	1801	CCATGGACTAAAGTCTGGTCTCATCTGAGCTCAGCTCGCTCACACAGCTCTAGGGGCTG	1860
QY	1861	CTCCTCTAACTCACAAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT	1920
Db	1861	CTCCTCTAACTCACAAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT	1920
QY	1921	GAGCAAAATAGCAAAAGCCTCTCTCAGCCACTCGCCTGAATCTACACTGGAAGCCAAC	1980
Db	1921	GAGCAAAATAGCAAAAGCCTCTCTCAGCCACTCGCCTGAATCTACACTGGAAGCCAAC	1980
QY	1981	TTGCTGGCACCCCGCTCCCAACCTCTTTCCTGGGTAGGAGAGGCTTAAAGATCACCC	2040
Db	1981	TTGCTGGCACCCCGCTCCCAACCTCTTTCCTGGGTAGGAGAGGCTTAAAGATCACCC	2040
QY	2041	TAAATTTACTCATCTCTTAGTGTCCCTCAATTTGGGCTCAGCAGCTCCCAAGCACCA	2100
Db	2041	TAAATTTACTCATCTCTTAGTGTCCCTCAATTTGGGCTCAGCAGCTCCCAAGCACCA	2100
QY	2101	ATTACAGGTACCCCTCTCTTCTTGCACTGTCGCCCAACTTGTCTCAATTTCCGAGATC	2160
Db	2101	ATTACAGGTACCCCTCTCTTCTTGCACTGTCGCCCAACTTGTCTCAATTTCCGAGATC	2160
QY	2161	TAACTCTCCCTTACGCTCTGCGAGAAATTTCTTTAGACCTCTACTAGCAAGCCGGTTG	2220
Db	2161	TAACTCTCCCTTACGCTCTGCGAGAAATTTCTTTAGACCTCTACTAGCAAGCCGGTTG	2220
QY	2221	CTCCTTGTGAGGAAATTTGTAGATCAATTTCAATTTTCCCTGGGCTGATCTTCT	2280
Db	2221	CTCCTTGTGAGGAAATTTGTAGATCAATTTCAATTTTCCCTGGGCTGATCTTCT	2280
QY	2281	CTCATCTTGCAACCCCAACCTCTGTGTAATAGATTTACCGCATTTACGGCTGCTTGTAA	2340

```

Db      2281  CTCACTCTGCAACCCCACTCTGTAATAGATTACCGCATTTACGGCTGCTCTGTAA 2340
QY      2341  GTGGGATGCTCTCCTAATGGAGAGTGTTTCATTTATTAATAGTTATTCACCTGAGTAT 2400
Db      2341  GTGGGATGCTCTCCTAATGGAGAGTGTTTCATTTATTAATAGTTATTCACCTGAGTAT 2400
QY      2401  GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456
Db      2401  GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456

RESULT 3
ADI28459
ID      ADI28459 standard; cDNA; 2456 BP.
XX      AC      ADI28459;
XX      DT      22-APR-2004 (first entry)
XX      DE      Human GPCR retinoic acid induced 3 (RAI-3) cDNA.
KW      Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW      antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW      broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
KW      chromosome 12p13-p12.3; s6.
XX      OS      Homo sapiens.
XX      FH      Key
XX      FT      Location/Qualifiers
XX      FT      replace(112,a)
XX      FT      /tag= a
XX      FT      /label= RAI-3-s1
XX      FT      /note= "located in 5' untranslated region"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      CDS
XX      FT      254..1327
XX      FT      /*tag= b
XX      FT      /product= "Human RAI3"
XX      FT      replace(364,t)
XX      FT      /*tag= c
XX      FT      /label= RAI-3-s2
XX      FT      /note= "exon 1, silent (Ala/Ala)"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      replace(51,t)
XX      FT      /*tag= d
XX      FT      /label= RAI-3-s3
XX      FT      /note= "exon 2, silent (Ile/Ile)"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      replace(53,t)
XX      FT      /*tag= e
XX      FT      /label= RAI-3-s4
XX      FT      /note= "exon 2, silent (Asp/Asp)"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      replace(605,g)
XX      FT      /*tag= f
XX      FT      /label= RAI-3-s6
XX      FT      /note= "exon 2, missense (Ser/Gly)"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      replace(797,g)
XX      FT      /*tag= g
XX      FT      /label= RAI-3-s5
XX      FT      /note= "exon 2, missense (Thr/Ala)"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      replace(1111,g)
XX      FT      /*tag= h
XX      FT      /label= RAI-3-s8
XX      FT      /note= "silent (Pro/Pro)"
XX      FT      /standard_name= "Single nucleotide polymorphism"
XX      FT      replace(1173,g)
XX      FT      /*tag= i
XX      FT      /label= RAI-3-s9
XX      FT      /note= "missense (Gln/Arg)"
XX      FT      /standard_name= "Single nucleotide polymorphism"

```

```

XX      WO2004001060-A2.
XX      PD      31-DEC-2003.
XX      PF      20-JUN-2003; 2003WO-US019255.
XX      PR      20-JUN-2002; 2002US-0390850P.
XX      PR      29-AUG-2002; 2002US-0407006P.
XX      PA      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX      PI      Whitney GS, Opitck G, Garulacan L, Ramanathan CS, McKinnon M;
XX      PI      Bennett KL, Barber LE, Cacace A, Teuchihashi Z;
XX      DR      WPI; 2004-090973/09.
XX      DR      P-PSDB; ADI28460.
XX      DR      GENBANK; NM_003979.
XX      PT      New nucleic acid molecule encoding a human G-protein coupled receptor,
XX      PT      RAI3, useful for preventing, treating or ameliorating chronic obstructive
XX      PT      pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
XX      PT      of COPD.
XX      PS      Claim 1; SEQ ID NO 2; 301pp; English.
XX      CC      The present sequence is that of cDNA encoding a human G-protein coupled
XX      CC      receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
XX      CC      Proteomics methods were used to isolate cigarette smoke-inducible
XX      CC      tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
XX      CC      identified as being tyrosine phosphorylated and/or as being
XX      CC      associated/complexed with tyrosine phosphorylated proteins only in those
XX      CC      cells that had been exposed to cigarette smoke. Since RAI-3 is primarily
XX      CC      expressed in lung tissue, and since cigarette smoke is a major causative
XX      CC      factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
XX      CC      novel cellular target for identifying modulators, e.g. agonists or
XX      CC      antagonists, useful for the treatment and/or prevention of COPD and
XX      CC      related disorders such as emphysema and chronic bronchitis. RAI-3
XX      CC      compounds, e.g. agonists and antagonists, especially antisense
XX      CC      compounds, can be used to treat COPD and other disorders and diseases
XX      CC      associated with regulation of NF-kB and/or its associated or interacting
XX      CC      signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
XX      CC      the RAI-3 gene are useful for determining COPD association in
XX      CC      individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
XX      CC      preventing, treating or ameliorating disorders related to aberrant GPCR
XX      CC      signaling or cell cycle regulation, pulmonary disorders, inflammatory
XX      CC      lung disorders, COPD, the underlying symptoms of COPD, COPD-related
XX      CC      disorders or conditions, autoimmune disorders, disorders related to
XX      CC      hyperimmune activity, inflammatory conditions, disorders related to
XX      CC      aberrant acute phase responses, hypercongenital conditions, birth
XX      CC      defects, necrotic lesions, wounds, organ transplant rejection, renal
XX      CC      diseases, ischaemia-reperfusion injury, heart disorders, disorders
XX      CC      related to aberrant signal transduction, proliferation disorders,
XX      CC      cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
XX      CC      ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
XX      CC      nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
XX      CC      rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
XX      CC      pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
XX      CC      colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
XX      CC      associated with aberrant cell adhesion, I-CAM function and/or regulation,
XX      CC      B-selectin function and/or regulation, or aberrant NF-kB function and/or
XX      CC      regulation (all claimed).
XX      SQ      Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match      100.0%; Score 2456; DB 12; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  ATAAAGCATGAAGTCCGTTGGAACTGGGAATAGCGGTGCTCTCCCTCGACCTCCCCC 60
Db      1  ATAAAGCATGAAGTCCGTTGGAACTGGGAATAGCGGTGCTCTCCCTCGACCTCCCCC 60

```

QY 61 TCCTTGTCCTCTGCTCAGCCCTCGCTCGTTCCCTCCCTCCGCGAGGGCCGCTTTATA 120
Db |||||
QY 61 TCCTTGTCCTCTGCTCAGCCCTCGCTCGTTCCCTCCCTCCGCGAGGGCCGCTTTATA 120
Db |||||
QY 121 ACAAATGCTCAGAGTGCAGAGGGCGGATAGCTGTCCAAAGGTCTCCGCCAGCACTGAGGAG 180
Db |||||
QY 121 ACNATGCTCAGAGTGCAGAGGGCGGATAGCTGTCCAAAGGTCTCCGCCAGCACTGAGGAG 180
Db |||||
QY 181 CTCGCTGCTGCCCTCTTCGCGCGCGGAAGAGCAGCAACAAAGTTCA CGGCCAA CGCCTTGGC 240
Db |||||
QY 241 ACTAGGTCAGNATGGCTACACAGTCCCTGATGTTGCGCGCAATGGCCTGAATCCAA 300
Db |||||
QY 241 ACTAGGTCAGNATGGCTACACAGTCCCTGATGTTGCGCGCAATGGCCTGAATCCAA 300
Db |||||
QY 301 GTACTACAGACTTGTGATAAGGCTGAAGCTTGGGGCATCGCTTAGAAACGGTGGCCAC 360
Db |||||
QY 301 GTACTACAGACTTGTGATAAGGCTGAAGCTTGGGGCATCGCTTAGAAACGGTGGCCAC 360
Db |||||
QY 361 AGCCGGGTTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420
Db |||||
QY 361 AGCCGGGTTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420
Db |||||
QY 421 GCAGGACTCCACAGCGGAAATGCTGCTACTCAGTTTCTCTCCTCGGTGTGTT 480
Db |||||
QY 421 GCAGGACTCCACAGCGGAAATGCTGCTACTCAGTTTCTCTCCTCGGTGTGTT 480
Db |||||
QY 481 GGGCATCTTTGGCTCACCTTCGCTTCATCATCGGACTGGACGGGAGCACAGGGCCAC 540
Db |||||
QY 481 GGGCATCTTTGGCTCACCTTCGCTTCATCATCGGACTGGACGGGAGCACAGGGCCAC 540
Db |||||
QY 541 AGCTTCTCTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Db |||||
QY 541 AGCTTCTCTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Db |||||
QY 601 TGTGAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTCCCTGTTGGTGATCTGGG 660
Db |||||
QY 601 TGTGAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTCCCTGTTGGTGATCTGGG 660
Db |||||
QY 661 TCTGGCCGTGGGCTTCAGCTAGTCAGGATGTTATCGCTATTGAATATATTGCTCTGAC 720
Db |||||
QY 661 TCTGGCCGTGGGCTTCAGCTAGTCAGGATGTTATCGCTATTGAATATATTGCTCTGAC 720
Db |||||
QY 721 CATGAATAGGACCAACAGTCAATCTCTTTCTGAGCTTCCGCTCTCTGCTGCAATGAAGA 780
Db |||||
QY 721 CATGAATAGGACCAACAGTCAATCTCTTTCTGAGCTTCCGCTCTCTGCTGCAATGAAGA 780
Db |||||
QY 781 CTTTGTCTCTGCTCAGCTAGTCTCTCTCTTGATGGGCTGACCTTCCTCATGTCTCTC 840
Db |||||
QY 781 CTTTGTCTCTGCTCAGCTAGTCTCTCTCTTGATGGGCTGACCTTCCTCATGTCTCTC 840
Db |||||
QY 841 CTTTACCTTCTGTGGTCTCTTCA CGGGCTGGAAGAGACATGGGGCCACATCTACTCTAC 900
Db |||||
QY 841 CTTTACCTTCTGTGGTCTCTTCA CGGGCTGGAAGAGACATGGGGCCACATCTACTCTAC 900
Db |||||
QY 901 GATGCTCTCTCCATTGGCCATCTGGGTGGCCTGGATCACCTGCTCATGCTTCTGACTT 960
Db |||||
QY 901 GATGCTCTCTCCATTGGCCATCTGGGTGGCCTGGATCACCTGCTCATGCTTCTGACTT 960
Db |||||
QY 961 TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCTGCAATGGCTGGGT 1020
Db |||||
QY 961 TGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCTGCAATGGCTGGGT 1020
Db |||||
QY 1021 CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAGAAAGCAACGAAACCCCATGGA 1080
Db |||||
QY 1021 CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAGAAAGCAACGAAACCCCATGGA 1080
Db |||||
QY 1081 TTATCCTGTTAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
Db |||||
QY 1081 TTATCCTGTTAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
Db |||||
QY 1141 GNAACAGGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGACAGGGGACACGCT 1200

Db |||||
QY 1141 GNAACAGGCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGACAGGGGACACGCT 1200
Db |||||
QY 1201 CTATGCCCCCTATTCCACACATTTTTCAGCTGCAGAACACAGCCTCCCCAAAAGGAATCTTC 1260
Db |||||
QY 1201 CTATGCCCCCTATTCCACACATTTTTCAGCTGCAGAACACAGCCTCCCCAAAAGGAATCTTC 1260
Db |||||
QY 1261 CATCCACAGGGCCCCACGTTGGCCGAGCCCTTACAAAAGACTATGAAGTAAAGAAAGAGGG 1320
Db |||||
QY 1261 CATCCACAGGGCCCCACGTTGGCCGAGCCCTTACAAAAGACTATGAAGTAAAGAAAGAGGG 1320
Db |||||
QY 1321 CAGCTTAACCTCTGCTGAGAGTGGGCAAAATGCGAGCCGGCGGAGACTAGCGGAGC 1380
Db |||||
QY 1321 CAGCTTAACCTCTGCTGAGAGTGGGCAAAATGCGAGCCGGCGGAGACTAGCGGAGC 1380
Db |||||
QY 1381 TCAAAAGGAGTGGGCGAAATCTTGAGTCTTCTGAGAAAATGTAACAAGACACTACGGGA 1440
Db |||||
QY 1381 TCAAAAGGAGTGGGCGAAATCTTGAGTCTTCTGAGAAAATGTAACAAGACACTACGGGA 1440
Db |||||
QY 1441 ACAGTTTGCCTCCCTCCAGGCTCAACCAAAATCTTCCATGCTGGGGCTGATGTGGCT 1500
Db |||||
QY 1441 ACAGTTTGCCTCCCTCCAGGCTCAACCAAAATCTTCCATGCTGGGGCTGATGTGGCT 1500
Db |||||
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTCTTTTGTCTCATCTTTGG 1560
Db |||||
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTCTTTTGTCTCATCTTTGG 1560
Db |||||
QY 1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTGTAGACCCCTTACTCTTTTGT 1620
Db |||||
QY 1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTGTAGACCCCTTACTCTTTTGT 1620
Db |||||
QY 1621 GTTTTTTGAACACAGATCTTGCTCTGTCAACCAGGCTTGAGTGCAAGTGTGGGATCACAG 1680
Db |||||
QY 1621 GTTTTTTGAACACAGATCTTGCTCTGTCAACCAGGCTTGAGTGCAAGTGTGGGATCACAG 1680
Db |||||
QY 1681 CCCAGTGCAGCTCCAGCACCTGTCTCAAGCAATCTCCCATCTCCATCTCCAAAGTG 1740
Db |||||
QY 1681 CCCAGTGCAGCTCCAGCACCTGTCTCAAGCAATCTCCCATCTCCATCTCCAAAGTG 1740
Db |||||
QY 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTTGCTGTTATTTT 1800
Db |||||
QY 1741 CTGGGATGACAGCGGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTTGCTGTTATTTT 1800
Db |||||
QY 1801 CCAATGCACTAAAGGTCTGCTCATCTGAGCTCAGCTGGCTCAACAGCTCTAGGGGCTG 1860
Db |||||
QY 1801 CCAATGCACTAAAGGTCTGCTCATCTGAGCTCAGCTGGCTCAACAGCTCTAGGGGCTG 1860
Db |||||
QY 1861 CTCCTCTAATCTCAGAGTGGGTTTCTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
Db |||||
QY 1861 CTCCTCTAATCTCAGAGTGGGTTTCTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
Db |||||
QY 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAAATCTACACTGGAAGCCAA 1980
Db |||||
QY 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAAATCTACACTGGAAGCCAA 1980
Db |||||
QY 1981 TTGCTGGCAACCCCGCTCCCAAACCTTCTTGCTGGGTAGGAGGCTTAAAGATCACCC 2040
Db |||||
QY 1981 TTGCTGGCAACCCCGCTCCCAAACCTTCTTGCTGGGTAGGAGGCTTAAAGATCACCC 2040
Db |||||
QY 2041 TAAATTTACTCATCTCTAGTGTGCTCAATTTGGGCTCAGCAGCTCCCGAGCACA 2100
Db |||||
QY 2041 TAAATTTACTCATCTCTAGTGTGCTCAATTTGGGCTCAGCAGCTCCCGAGCACA 2100
Db |||||
QY 2101 ATTTACAGGTCAACCCCTCTCTTTGCACTGTCCCAAACTTTGCTCAATTTCCGAGATC 2160
Db |||||
QY 2161 TAACTCTCCCTACGCTCTGCCAGAAATCTTTTCAAGCTCTACTAGCAACAGCCGGTTG 2220
Db |||||
QY 2161 TAACTCTCCCTACGCTCTGCCAGAAATCTTTTCAAGCTCTACTAGCAACAGCCGGTTG 2220
Db |||||
QY 2221 CTCCTTGTTCAGGAGAAATTTGTAGATCAATCTCAATTTCAAATTTCTGGGCTGATCTTCT 2280
Db |||||

Db 2221 CTCTTTGTCAGGAGATTTGTAGATCATTTCTCAATTTCTCGGGCTGATCTTCT 2280
Qy 2281 CTATCTTGACCCACCTCTGTAATAGATTTACCGCATTTACGGCTGCTATCTGTAA 2340
Db 2281 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2340
Qy 2341 GTGGGATGCTCTCTTAATAGGAGAGTGTTCATTTGATATATAAGTTATTCACCTGAGTAT 2400
Db 2341 GTGGGATGCTCTCTTAATAGGAGAGTGTTCATTTGATATATAAGTTATTCACCTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGGTGGCCACCTCTTTCATGGTGGTGGAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGGCCACCTCTTTCATGGTGGTGGAGCAAAAAA 2456

RESULT 4

ADR48221 standard; cDNA; 2456 BP.

XX AC ADR48221;
XX 18-NOV-2004 (first entry)
XX Human retinoic acid induced 3 (RAI3) encoding cDNA SEQ.9.
DE pancreatic cancer-associated transcript; pancreatic cancer; human;
KW cytostatic; gene therapy; protein therapy; retinoic acid induced 3; RAI3;
KW gene, ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 254..1327
FT /tag= a
FT /product= "retinoic acid induced 3 (RAI3)"
XX WO2004074510-A1.
PN
XX
XX 02-SEP-2004.
XX
PF 18-FEB-2004; 2004WO-AU000194.
XX
PR 18-FEB-2003; 2003AU-00900747.
XX
XX (GARV-) GARVAN INST MEDICAL RES.
XX
XX Biankin A, Segara D, Henshall S, Sutherland R;
PI
XX
XX WPI; 2004-635591/61.
DR P-PSDB; ADR48222.
XX

PT Detecting pancreatic cancer-associated transcript in a biological sample,
PT useful for diagnosing or treating the disease, comprises contacting the
PT sample with a polynucleotide that selectively hybridizes to a specific
PT sequence.

XX Claim 7; SEQ ID NO 9; 263pp; English.

XX The present invention describes a method for detecting a pancreatic
XX cancer-associated transcript in a biological sample. The method comprises
XX contacting the biological sample with a polynucleotide that selectively
XX hybridizes to a sequence at least 80% identical to a sequence as shown in
XX any one of Tables 3 to 25 in the specification or having the GenBank
XX Accession Number AF279145. Also described: (1) diagnosing pancreatic
XX cancer in a human or animal subject being tested, determining the
XX likelihood that a subject having a pancreatic cancer will survive, or
XX determining the suitability of a subject having a pancreatic cancer for
XX surgical resection therapy; (2) detecting a pancreatic cancer-associated
XX polypeptide in a biological sample; (3) determining the likelihood that a
XX subject having a pancreatic cancer will survive; and (4) monitoring the
XX efficacy of a therapeutic treatment of pancreatic cancer. A pancreatic
XX cancer-associated transcript has cytostatic activity, and can be used in
XX gene and protein therapy. A pancreatic cancer-associated transcript

CC polynucleotide, a vector comprising the polynucleotide, an isolated
CC polypeptide or an antibody that binds to the isolated polypeptide can be
CC used for diagnosing or prognosing pancreatic cancer or for preparing a
CC medicament for the treatment of pancreatic cancer. The prognostic or
CC diagnostic methods are useful for the early detection of pancreatic
CC cancer or its metastases, and for monitoring the progress of disease such
CC as during remission or following surgery or chemotherapy. The present
CC sequence encodes human retinoic acid induced 3 (RAI3), which is used in
CC the exemplification of the present invention.

XX SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 2456; DB 13; Length 2456;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATAACAGCATGAAGTCCGCTGGAACTGGAAATAGCGGTGCTCTCCCTCGAGGGCGCTTTATA 120

Db 1 ATAACAGCATGAAGTCCGCTGGAACTGGAAATAGCGGTGCTCTCCCTCGAGGGCGCTTTATA 120

Qy 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGGCGCTTTATA 120

Db 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGGCGCTTTATA 120

Qy 121 ACAACTGCTCAGAGTGGAGGGCGGATAGCTGTCCAAGGTCTCCCCAGCACTGAGGAG 180

Db 121 ACAACTGCTCAGAGTGGAGGGCGGATAGCTGTCCAAGGTCTCCCCAGCACTGAGGAG 180

Qy 181 CTCGCTCTGCTCCCTCTTGGCGCGGGAAGCAGACACAAAGTTCACGGCCAAACGCTTGGC 240

Db 181 CTCGCTCTGCTCCCTCTTGGCGCGGGAAGCAGACACAAAGTTCACGGCCAAACGCTTGGC 240

Qy 241 ACTAGGGTCCAGATGGCTTACACAGTCCCTGATGCTCCGCAATGGCTGAAATCCAA 300

Db 241 ACTAGGGTCCAGATGGCTTACACAGTCCCTGATGCTCCGCAATGGCTGAAATCCAA 300

Qy 301 GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGCGATCGTCTAGAAACGGTGGCCAC 360

Db 301 GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGCGATCGTCTAGAAACGGTGGCCAC 360

Qy 361 AGCCGGGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420

Db 361 AGCCGGGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420

Qy 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTGGGTGGTT 480

Db 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTGGGTGGTT 480

Qy 481 GGGCATCTTTGGCTCACCTTCGCTTCATCATCGGACTGGAGGAGGAGGAGGAGGAGGAG 540

Db 481 GGGCATCTTTGGCTCACCTTCGCTTCATCATCGGACTGGAGGAGGAGGAGGAGGAGGAG 540

Qy 541 AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTCTGCTGCTCATGC 600

Db 541 AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTCTGCTGCTCATGC 600

Qy 601 TGTCACTGACCAAGCTCGTCCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

Db 601 TGTCACTGACCAAGCTCGTCCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

Qy 661 TCTGGCCGTGGGCTTCAGCTAGTCCAGGATGTTATCGTATTGAATATATATTCCTCTGAC 720

Db 661 TCTGGCCGTGGGCTTCAGCTAGTCCAGGATGTTATCGTATTGAATATATATTCCTCTGAC 720

Qy 721 CATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTCCGCTCTCTGCGCAATGAAGA 780

Db 721 CATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTCCGCTCTCTGCGCAATGAAGA 780

Qy 781 CTTTGTCTCTCTGCTCACCTAGCTCTCTCTTGTATGGCGCTGACCTTCTCATGTCCTC 840

Db 781 CTTTGTCTCTCTGCTCACCTAGCTCTCTCTTGTATGGCGCTGACCTTCTCATGTCCTC 840

Qy 841 CTTTACCTTCTGTGGTTCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900

Db 841 CTTACCTTTCTGGTTCCTTCAOGGCTGGAAGACATGGGGCCACATCTACCTCAC 900
Qy 901 GATGTCCTCTCATTTGCCATCTGGGTGGCTGGATCACCTGCTCATCTCTCTGACTT 960
Db 901 GATGTCCTCTCATTTGCCATCTGGGTGGCTGGATCACCTGCTCATCTCTGACTT 960
Qy 961 TGAACGAGGTGGATGACACCATCTCAGCTCCGCTTTGGCTGGCCAAATGGCTGGGTGT 1020
Db 961 TGAACGAGGTGGATGACACCATCTCAGCTCCGCTTTGGCTGGCCAAATGGCTGGGTGT 1020
Qy 1021 CTTGTTGGCTTATGTTAGTCCGAGTTTGGCTGGCTCAAAAGCAACGAAACCCCATGA 1080
Db 1021 CTTGTTGGCTTATGTTAGTCCGAGTTTGGCTGGCTCAAAAGCAACGAAACCCCATGA 1080
Qy 1081 TTATCTGTTGAGGATGCTTTCTGTAAGCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
Db 1081 TTATCTGTTGAGGATGCTTTCTGTAAGCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
Qy 1141 GAAACAGAGCTACTCTCAAGAGGAATCACTCAAGGTTTGAAGAGACAGGGAACGCT 1200
Db 1141 GAAACAGAGCTACTCTCAAGAGGAATCACTCAAGGTTTGAAGAGACAGGGAACGCT 1200
Qy 1201 CTATGCCCTTATTCACACATTTTCAGCTGAGACACGCTCCCAAGAAAGATTTCT 1260
Db 1201 CTATGCCCTTATTCACACATTTTCAGCTGAGACACGCTCCCAAGAAAGATTTCT 1260
Qy 1261 CATCCACGGGCCACGCTTGGCCGAGCCTTACAAAGACTATGAAGTAAAGAAAGAGG 1320
Db 1261 CATCCACGGGCCACGCTTGGCCGAGCCTTACAAAGACTATGAAGTAAAGAAAGAGG 1320
Qy 1321 CAGCTAACTCTGCTCTGAAGAGTGGACAAATGACAGCCGGGGCGAGATCTAGCGGAGC 1380
Db 1321 CAGCTAACTCTGCTCTGAAGAGTGGACAAATGACAGCCGGGGCGAGATCTAGCGGAGC 1380
Qy 1381 TCNAAGGATGTTGGGGAATCTTGAGTCTCTGAGAAAACGTGTAACAGACACTAGCGA 1440
Db 1381 TCNAAGGATGTTGGGGAATCTTGAGTCTCTGAGAAAACGTGTAACAGACACTAGCGA 1440
Qy 1441 ACAGTTTGCTCCTCCACGCTCAACCAAACTTCTTCATGCTGGGGCTGATGGGCT 1500
Db 1441 ACAGTTTGCTCCTCCACGCTCAACCAAACTTCTTCATGCTGGGGCTGATGGGCT 1500
Qy 1501 AGTAAGCTCCAGTCTTGAAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTCTGG 1560
Db 1501 AGTAAGCTCCAGTCTTGAAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTCTGG 1560
Qy 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
Db 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
Qy 1621 GTTTTTTGAACAGGATCTTGCTCTGTCAACCCAGGCTTGAGTGCAGTGGTGCATCACAG 1680
Db 1621 GTTTTTTGAACAGGATCTTGCTCTGTCAACCCAGGCTTGAGTGCAGTGGTGCATCACAG 1680
Qy 1681 CCAGTGCAGCTTCGACCACTGTGTCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db 1681 CCAGTGCAGCTTCGACCACTGTGTCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Qy 1741 CTGGGATGACGGCGTGGCCACAGCTCCAGCCTAGGCCCTTAATCTTGCTGTTATTT 1800
Db 1741 CTGGGATGACGGCGTGGCCACAGCTCCAGCCTAGGCCCTTAATCTTGCTGTTATTT 1800
Qy 1801 CCATGACTAAAGTCTGCTCATCTCAGCTCAGCTGGCTCACACAGCTCTAGGGCCCTG 1860
Db 1801 CCATGACTAAAGTCTGCTCATCTCAGCTCAGCTGGCTCACACAGCTCTAGGGCCCTG 1860
Qy 1861 CTCCTCTAACTCAGTGGGTTTTGTAGGCTCTGTGGCCACAGACAGCTTCATATCT 1920
Db 1861 CTCCTCTAACTCAGTGGGTTTTGTAGGCTCTGTGGCCACAGACAGCTTCATATCT 1920
Qy 1921 GAGCAAAATAGCAAAAGCTCTCTCAGCCACTGGCCTGAAATCTACACTGGAAGCAAC 1980

Db 1921 GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAGCAAC 1980
Qy 1981 TTGCTGGCAACCCCGCTCCCAACCCCTTCTTGCTGGGTAGGAGGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCAACCCCGCTCCCAACCCCTTCTTGCTGGGTAGGAGGGCTAAAGATCACCC 2040
Qy 2041 TAAATTTACTCATCTCTTAGTGTGCTCATATGGGCTCAGAGCTCCCGACCA 2100
Db 2041 TAAATTTACTCATCTCTTAGTGTGCTCATATGGGCTCAGAGCTCCCGACCA 2100
Qy 2101 ATTACAGGTCAACCCCTCTCTTCTTGCACTGCCAAACTTGTCTCAATTCGAGATC 2160
Db 2101 ATTACAGGTCAACCCCTCTCTTCTTGCACTGCCAAACTTGTCTCAATTCGAGATC 2160
Qy 2161 TAATCTCCCTTACGCTCTGCCAGGAATTTCTTTCAGACCTCACTAGCAAAAGCCGGTTG 2220
Db 2161 TAATCTCCCTTACGCTCTGCCAGGAATTTCTTTCAGACCTCACTAGCAAAAGCCGGTTG 2220
Qy 2221 CTCCTTGTGAGAGAAATTTGTAGATCATTTCTCACTTCAATTCCTGGGCTGATCTTCT 2280
Db 2221 CTCCTTGTGAGAGAAATTTGTAGATCATTTCTCACTTCAATTCCTGGGCTGATCTTCT 2280
Qy 2281 CTCATCTTGCACCCCAACCTCTGTAATAGATTTACCGCATTTACGGCTGCATTTCTGTAA 2340
Db 2281 CTCATCTTGCACCCCAACCTCTGTAATAGATTTACCGCATTTACGGCTGCATTTCTGTAA 2340
Qy 2341 GTGGGATGCTCTCTTAATGGAGAGTGTTCATTTGATTAATAGTTATTCACCTGAGTAT 2400
Db 2341 GTGGGATGCTCTCTTAATGGAGAGTGTTCATTTGATTAATAGTTATTCACCTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456

RESULT 5
ACN39388
ID ACN39388 standard; cDNA; 2456 BP.
XX
AC ACN39388;
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA226771, SEQ ID NO:3498.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
XX gene therapy; cytostatic; gene; ss.
XX Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX P-PSDB; ABM81354.
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or

PT	prostate cancer or tumor.	
XX	Claim 1; SEQ ID NO 3498; 7273pp; English.	
PS	The invention relates to human tumour-associated antigenic target (TAT)	
XX	polypeptides, and their related nucleic acids. The TAT polypeptides are	
CC	overexpressed in cancer tissues compared to normal tissues, and may thus	
CC	serve as effective targets for the diagnosis and treatment of cancer in	
CC	mammals. The invention also relates to nucleic acid and polypeptide	
CC	sequences at least 80% identical to the TAT nucleic acids and	
CC	polypeptides; expression vectors and host cells comprising a TAT nucleic	
CC	acid; an antibody specific for a TAT polypeptide; a peptide or organic	
CC	molecule which binds to a TAT polypeptide; fusion proteins comprising a	
CC	TAT polypeptide; and methods and compositions for the treatment or	
CC	diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,	
CC	antibodies, antagonists, binding molecules and compositions are useful	
CC	for diagnosing or treating a cell proliferative disorder associated with	
CC	increased TAT expression, particularly cancers, such as breast cancer,	
CC	colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder	
CC	cancer, pancreatic cancer, cervical cancer, cancers of the central	
CC	nervous system, melanoma and leukaemia. TAT nucleic acids may further be	
CC	used as hybridisation probes, in chromosome and gene mapping, in	
CC	chromosome identification and in gene therapy. The present sequence	
CC	represents a TAT nucleic acid of the invention	
XX		
SQL	Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 2456; DB 13; Length 2456;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATAAACAGCATGAAGTGCCTGGAACTCGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCCC	60
DB	1 ATAAACAGCATGAAGTGCCTGGAACTCGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCCC	60
QY	61 TCCTTGTCCCTCTGCTCAGCCCTCGCTCGTTCCTCCCTCGGCGAGGGCGGCTTTATA	120
DB	61 TCCTTGTCCCTCTGCTCAGCCCTCGCTCGTTCCTCCCTCGGCGAGGGCGGCTTTATA	120
QY	121 ACAACTCTCAGAGTGCAGGGGGGATAGCTGTCCAAGGTCTCCCCAGCACCTAGGAG	180
DB	121 ACAACTCTCAGAGTGCAGGGGGGATAGCTGTCCAAGGTCTCCCCAGCACCTAGGAG	180
QY	181 CTCGCTCTGCTCCCTCTTGGCGGGGGAAGCAGCACCAAGTTACGCGCCAACTGGC	240
DB	181 CTCGCTCTGCTCCCTCTTGGCGGGGGAAGCAGCACCAAGTTACGCGCCAACTGGC	240
QY	241 ACTAGGTCTCAGAAATGGCTCAACAGTCCCTGATGGTTGGCGCAATGGCTGAAATCCAA	300
DB	241 ACTAGGTCTCAGAAATGGCTCAACAGTCCCTGATGGTTGGCGCAATGGCTGAAATCCAA	300
QY	301 GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360
DB	301 GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360
QY	361 AGCGGGGTTGTGACCTCGTGGGCTTTATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420
DB	361 AGCGGGGTTGTGACCTCGTGGGCTTTATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420
QY	421 GCAGGACTCCAAAGGCGAAATATGCTGCTTACTCAGTTTCTCTCTCTCTCTCTCTCTCT	480
DB	421 GCAGGACTCCAAAGGCGAAATATGCTGCTTACTCAGTTTCTCTCTCTCTCTCTCTCTCT	480
QY	481 GGGCATCTTGGGCTCAGCTTCGCTTCATCATCGGACTGGAACGGGAGCAGGGGCCAC	540
DB	481 GGGCATCTTGGGCTCAGCTTCGCTTCATCATCGGACTGGAACGGGAGCAGGGGCCAC	540
QY	541 ACGCTTCT	600
DB	541 ACGCTTCT	600
QY	601 TGTGAGTCTGACCAAGCTCGTTCGGGGGAGGAAGCCCTTTTCCCTGTGTGTGATTTCTGGG	660
DB	601 TGTGAGTCTGACCAAGCTCGTTCGGGGGAGGAAGCCCTTTTCCCTGTGTGTGATTTCTGGG	660
QY	661 TCTGGCGCTGGGCTTCCAGCTAGTCCAGGATGTTATCGCTATTGAAATATATTCTCTGAC	720
DB	661 TCTGGCGCTGGGCTTCCAGCTAGTCCAGGATGTTATCGCTATTGAAATATATTCTCTGAC	720
QY	721 CATGAATAGGACCAACGTCCTTTCTGAGCTTTTCCGCTCTCGTGCGAATGAAGA	780
DB	721 CATGAATAGGACCAACGTCCTTTCTGAGCTTTTCCGCTCTCGTGCGAATGAAGA	780
QY	781 CTTTGTCT	840
DB	781 CTTTGTCT	840
QY	841 CTTTCACTCTCTGCTGGTTCCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC	900
DB	841 CTTTCACTCTCTGCTGGTTCCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC	900
QY	901 GATGCTCTCTCTCATTTGGCTGCTGGATCAGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT	960
DB	901 GATGCTCTCTCTCATTTGGCTGCTGGATCAGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT	960
QY	961 TGACCCGAGGTGGATGACACATCTCTCAGCTCCGCTTGGCTGGAATGGCTGGGTGTT	1020
DB	961 TGACCCGAGGTGGATGACACATCTCTCAGCTCCGCTTGGCTGGAATGGCTGGGTGTT	1020
QY	1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA	1080
DB	1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA	1080
QY	1081 TTAATCTCTGTGAGGATGCTTTTCTGTAACCTCAACTCGTGAAGAAGAGCTATCGTGTGGA	1140
DB	1081 TTAATCTCTGTGAGGATGCTTTTCTGTAACCTCAACTCGTGAAGAAGAGCTATCGTGTGGA	1140
QY	1141 GAACAGGCTACTCTCAAGAGGAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCT	1200
DB	1141 GAACAGGCTACTCTCAAGAGGAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCT	1200
QY	1201 CTATGCCCCCTATTTCCACACATTTTTCAGCTGCAAGAACAGGCTTCCCAAGAGAAATTC	1260
DB	1201 CTATGCCCCCTATTTCCACACATTTTTCAGCTGCAAGAACAGGCTTCCCAAGAGAAATTC	1260
QY	1261 CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAGAGCTATGAAGTAAAGAAAGAGGG	1320
DB	1261 CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAGAGCTATGAAGTAAAGAAAGAGGG	1320
QY	1321 CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGCACGCGGGGGGAGAGCTACAGGGGAGC	1380
DB	1321 CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGCACGCGGGGGGAGAGCTACAGGGGAGC	1380
QY	1381 TCAAGGGATGTGGGCGAAATCTTTGAGTCTTCTGAGAAACTCTGACAGACACTACGGGA	1440
DB	1381 TCAAGGGATGTGGGCGAAATCTTTGAGTCTTCTGAGAAACTCTGACAGACACTACGGGA	1440
QY	1441 ACAGTTTGGCTCTCTCCAGCCCTCAACCAATTTCTTCCATGCTGGGGCTGATGTGGGCT	1500
DB	1441 ACAGTTTGGCTCTCTCCAGCCCTCAACCAATTTCTTCCATGCTGGGGCTGATGTGGGCT	1500
QY	1501 AGTAAGACTCCAGTCTTATAGGGCGCTGTAGTATTTTTTTTTTTTTTTTTTTTTTTTTT	1560
DB	1501 AGTAAGACTCCAGTCTTATAGGGCGCTGTAGTATTTTTTTTTTTTTTTTTTTTTTTTTT	1560
QY	1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTATAGCCCTTACTCTTTTGT	1620
DB	1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTATAGCCCTTACTCTTTTGT	1620
QY	16	

QY 1741 CTGGGATGACGGCGTGAGCCACAGCTCCAGCCTAGAGCCCTTAATCTTGTCTGTTATTTT 1800
Db 1741 CTGGGATGACGGCGTGAGCCACAGCTCCAGCCTAGAGCCCTTAATCTTGTCTGTTATTTT 1800
QY 1801 CCATGGACTAAAGGCTGGTCTATCTGAGCTCAGCTCAGCTGGCTCACAGCTCTAGGGCCCTG 1860
Db 1801 CCATGGACTAAAGGCTGGTCTATCTGAGCTCAGCTCAGCTGGCTCACAGCTCTAGGGCCCTG 1860
QY 1861 CTCCTCTAACTCACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCTATATCT 1920
Db 1861 CTCCTCTAACTCACAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCTATATCT 1920
QY 1921 GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAATCTACAGTGAAGCCAAAC 1980
Db 1921 GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCCTGAATCTACAGTGAAGCCAAAC 1980
QY 1981 TTGCTGGACCCCGCTCCCAACCTTCTTGGCTGGGTAGGAGGCTTAAAGATCACCC 2040
Db 1981 TTGCTGGACCCCGCTCCCAACCTTCTTGGCTGGGTAGGAGGCTTAAAGATCACCC 2040
QY 2041 TAAATTTACTCATCTCTAGTGGCTGCCTCAATTTGGGCTCAGCAGCTCCCGCAGCACCA 2100
Db 2041 TAAATTTACTCATCTCTAGTGGCTGCCTCAATTTGGGCTCAGCAGCTCCCGCAGCACCA 2100
QY 2101 ATTACAGTCAACCCCTCTCTTTTGCACTGTCCCAAACTTGTCTCAATTTCCGAGATC 2160
Db 2101 ATTACAGTCAACCCCTCTCTTTTGCACTGTCCCAAACTTGTCTCAATTTCCGAGATC 2160
QY 2161 TAAATTTACTCATCTCTAGTGGCTGCCTCAATTTGGGCTCAGCAGCTCCCGCAGCACCA 2220
Db 2161 TAAATTTACTCATCTCTAGTGGCTGCCTCAATTTGGGCTCAGCAGCTCCCGCAGCACCA 2220
QY 2221 CTCCTTGTGAGGAAATTTGTAGATCATCTCTCAATTTCTGGGCTGATATCTCT 2280
Db 2221 CTCCTTGTGAGGAAATTTGTAGATCATCTCTCAATTTCTGGGCTGATATCTCT 2280
QY 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTACCGCATTTACGGCTGCAATCTGTAA 2340
Db 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTACCGCATTTACGGCTGCAATCTGTAA 2340
QY 2341 GTGGGATGGTCTCTAATGGAGGAGTGTTCATTGTATATAAGTTATTCACCTGAGTAT 2400
Db 2341 GTGGGATGGTCTCTAATGGAGGAGTGTTCATTGTATATAAGTTATTCACCTGAGTAT 2400
QY 2401 GCAATAAAGATGTGGTGGCACTCTTTTCATGGTGGGAGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGGCACTCTTTTCATGGTGGGAGCAGCAAAAAA 2456

RESULT 6

ADU06126

ID ADU06126 standard; DNA; 2456 BP.

XX AC ADU06126;

XX AC ADU06126;

XX AC ADU06126;

DT 27-JAN-2005 (first entry)

DE DE Novel bronchial cancer-associated human gene SeqID350.

XX DE Novel bronchial cancer-associated human gene SeqID350.

KW bronchial cancer; cytostatic; tumour-associated protein;

XX KW bronchial cancer; cytostatic; tumour-associated protein;

OS Homo sapiens.

XX OS Homo sapiens.

PN DE10316701-Al.

XX PN DE10316701-Al.

PD 04-NOV-2004.

XX PD 04-NOV-2004.

PF 09-APR-2003; 2003DE-01016701.

XX PF 09-APR-2003; 2003DE-01016701.

PR 09-APR-2003; 2003DE-01016701.

XX PR 09-APR-2003; 2003DE-01016701.

(HINZ/) HINZMANN B.

(HERM/) HERMANN K.

(CAST/) HEIDEN CASTANOS-VELEZ E.

PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;

PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;

XX WPI; 2004-786403/78.

DR P-PSDB; ADU06613.

XX

XX

PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial

PT cancer and in screening for therapeutic and diagnostic agents.

XX

XX

PS Claim 1; SEQ ID NO 350; 1381pp; German.

XX

XX

CC This invention relates to a novel isolated nucleic acid associated with

CC bronchial cancer comprising 489 defined sequences given in the

CC specification. The invention may be useful for the production of

CC compounds with a cytostatic activity through the inhibition of expression

CC or activity of tumour-associated proteins. The novel DNA sequences and

CC the proteins/peptides encoded by them are used for detecting bronchial

CC cancer or determining the risk of developing it and to screen for

CC specific binding partners of the DNA or protein sequences, where the

CC binding partners are potentially useful as agents for treating or

CC diagnosing bronchial cancer. The DNA or protein sequences can also be

CC used for prognosis, detection of metastases and for secondary treatment

CC (of tumours that have been stabilised or are no longer detectable).

CC Detecting abnormal expression of the DNA sequences provides early

CC diagnosis of bronchial cancers. The present sequence is that of a novel

CC bronchial cancer-associated human gene sequence of the invention.

XX

SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 2456; DB 13; Length 2456;

Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAACAGCATGAAGTGGCGTGAACATGGAATAGGGGTGTCTCTCCCTCGACCTCCCC 60

Db 1 ATAAACAGCATGAAGTGGCGTGAACATGGAATAGGGGTGTCTCTCCCTCGACCTCCCC 60

QY 61 TCCTTGTCCCTCTGCTCACCCTCGCTCCCTCCCTCCCTCCCGGAGGCGCCCTTTATA 120

Db 61 TCCTTGTCCCTCTGCTCACCCTCGCTCCCTCCCTCCCTCCCGGAGGCGCCCTTTATA 120

QY 121 ACAACTGTCTCAGAGTGCAGAGGCGGGATAGCTGTCTCAAGGTCTCCCGCAGCACTGAGGAG 180

Db 121 ACAACTGTCTCAGAGTGCAGAGGCGGGATAGCTGTCTCAAGGTCTCCCGCAGCACTGAGGAG 180

QY 181 CTCGCTGTGCTGCCCTCTTTCGCGCGGGAAGAGCAGCACCAGTTTCACGGCCAAAGCCCTTGGC 240

Db 181 CTCGCTGTGCTGCCCTCTTTCGCGCGGGAAGAGCAGCACCAGTTTCACGGCCAAAGCCCTTGGC 240

QY 241 ACTAGGTTCCAGATGGCTTACACAGTCCCTCATGTTGGCGCAATGGCTGAAATCCAA 300

Db 241 ACTAGGTTCCAGATGGCTTACACAGTCCCTCATGTTGGCGCAATGGCTGAAATCCAA 300

QY 301 GTACTACAGACTTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC 360

Db 301 GTACTACAGACTTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC 360

QY 361 AGCGGGGTTGTGACCTCGGTGGCCCTTATGCTCACTCTCCGATCTCGTCTGCAAGGT 420

Db 361 AGCGGGGTTGTGACCTCGGTGGCCCTTATGCTCACTCTCCGATCTCGTCTGCAAGGT 420

QY 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTTCTCTCTGGTGTGTT 480

Db 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTTCTCTCTGGTGTGTT 480

QY 481 GGGCATCTTTGGCCTCACCTTCGCCCTTCATATCGGACTGGAACAGGCGCCAC 540

Db 481 GGGCATCTTTGGCCTCACCTTCGCCCTTCATATCGGACTGGAACAGGCGCCAC 540

```
QY 541 ACGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Db |||||
QY 541 ACGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC 600
Db |||||
QY 601 TGTCAGTCGACCAAGCTCTGTCGGGGAGGAAGCCCTTTTCCCTGTTGTTGATTCCTGGG 660
Db |||||
QY 601 TGTCAGTCGACCAAGCTCTGTCGGGGAGGAAGCCCTTTTCCCTGTTGTTGATTCCTGGG 660
Db |||||
QY 661 TCTGGCCGTGGGCTTCAGCCTAGTCAGAGTATTCGCTATTGAATATATTGCTCTGAC 720
Db |||||
QY 661 TCTGGCCGTGGGCTTCAGCCTAGTCAGAGTATTCGCTATTGAATATATTGCTCTGAC 720
Db |||||
QY 721 CATGAATAGAACCAACGTCATGCTCTTTTCTGAGCTTTCCGCTCCCTCGTCGCAATGAAGA 780
Db |||||
QY 721 CATGAATAGAACCAACGTCATGCTCTTTTCTGAGCTTTCCGCTCCCTCGTCGCAATGAAGA 780
Db |||||
QY 781 CTTTGTCTCTCTGCTCACCTACCTCTCTCTTCTTGTGATGGGCTGACCTTCTCATGTCCTC 840
Db |||||
QY 841 CTTACCTTCTGCTGCTTCTTACGGGCTGGAAGACATGGGGCCCAACATCTACCTCAC 900
Db |||||
QY 841 CTTACCTTCTGCTGCTTCTTACGGGCTGGAAGACATGGGGCCCAACATCTACCTCAC 900
Db |||||
QY 901 GATGCTCCTCTCCATTTGCCATCTGCGGTGGCTGGATCACCCTGCTCATGCTTCTGACTT 960
Db |||||
QY 901 GATGCTCCTCTCCATTTGCCATCTGCGGTGGCTGGATCACCCTGCTCATGCTTCTGACTT 960
Db |||||
QY 961 TGACCCAGGTGGGATGAACCAATCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGGTGTT 1020
Db |||||
QY 961 TGACCCAGGTGGGATGAACCAATCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGGTGTT 1020
Db |||||
QY 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAACCCCATGA 1080
Db |||||
QY 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAACCCCATGA 1080
Db |||||
QY 1081 TTATCCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAAGAGCTATGGTGTGA 1140
Db |||||
QY 1081 TTATCCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAAGAGCTATGGTGTGA 1140
Db |||||
QY 1141 GAAACAGAGCTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGACAGGGGACACGCT 1200
Db |||||
QY 1141 GAAACAGAGCTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGACAGGGGACACGCT 1200
Db |||||
QY 1201 CTATGCCCCCTATTCCACATTTTTCAGCTGAGAACCCCTCCCAAGAAAGAAATTCCTC 1260
Db |||||
QY 1201 CTATGCCCCCTATTCCACATTTTTCAGCTGAGAACCCCTCCCAAGAAAGAAATTCCTC 1260
Db |||||
QY 1261 CATCCACAGGCGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG 1320
Db |||||
QY 1261 CATCCACAGGCGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG 1320
Db |||||
QY 1321 CAGCTAACTCTGTCCTGAAGAGTGGGACAAATCAGCCGCGGCGGAGATCTAGCGGAGC 1380
Db |||||
QY 1321 CAGCTAACTCTGTCCTGAAGAGTGGGACAAATCAGCCGCGGCGGAGATCTAGCGGAGC 1380
Db |||||
QY 1381 TCAAGGAGTCTGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAGCACTACGGGA 1440
Db |||||
QY 1381 TCAAGGAGTCTGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAGCACTACGGGA 1440
Db |||||
QY 1441 ACAGTTTGCCTCCCTCCAGCCTCAACCAATTTCTTCCATGCTGGGCTGATGGGCT 1500
Db |||||
QY 1441 ACAGTTTGCCTCCCTCCAGCCTCAACCAATTTCTTCCATGCTGGGCTGATGGGCT 1500
Db |||||
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTCTTGG 1560
Db |||||
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTCTTGG 1560
Db |||||
QY 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCCCTACTCTTTTGT 1620
Db |||||
QY 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCCCTACTCTTTTGT 1620
Db |||||
QY 1621 GTTTTTTGAACAGGATCTTGCTCTGTCAACCCAGGCTTGAGTCAGTGGGATCACAG 1680
```

```
Db |||||
QY 1621 GTTTTTTGAACAGGATCTTGCTCTGTCAACCCAGGCTTGAGTCAGTGGGATCACAG 1680
Db |||||
QY 1681 CCAGTGCAGCCTCGAACCACTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db |||||
QY 1681 CCAGTGCAGCCTCGAACCACTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db |||||
QY 1741 CTGGGATGACAGGCTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTCTGTTATTTT 1800
Db |||||
QY 1741 CTGGGATGACAGGCTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTCTGTTATTTT 1800
Db |||||
QY 1801 CCATGACTTAAAGGTCTGGTCACTCTGAGCTCAGCTGGCTCAGACAGCTCTAGGGGCTG 1860
Db |||||
QY 1801 CCATGACTTAAAGGTCTGGTCACTCTGAGCTCAGCTGGCTCAGACAGCTCTAGGGGCTG 1860
Db |||||
QY 1861 CTCCTCTAACTACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGAGCTCGATATCT 1920
Db |||||
QY 1861 CTCCTCTAACTACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGAGCTCGATATCT 1920
Db |||||
QY 1921 GAGCAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAC 1980
Db |||||
QY 1921 GAGCAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAC 1980
Db |||||
QY 1981 TTGCTGGCACCCCGCTCCCCAAACCTTCTTGCTGGTAGGAGAGGCTAAAGATCACCC 2040
Db |||||
QY 1981 TTGCTGGCACCCCGCTCCCCAAACCTTCTTGCTGGTAGGAGAGGCTAAAGATCACCC 2040
Db |||||
QY 2041 TAAATTTACTCATCTCTTAGTGTGCTCACAATGGGGCTCAGAGCTCCCGAGACCA 2100
Db |||||
QY 2041 TAAATTTACTCATCTCTTAGTGTGCTCACAATGGGGCTCAGAGCTCCCGAGACCA 2100
Db |||||
QY 2101 ATTACAGGTGACCCCTCTCTTCTGCACTGCCCCAACTTGTCAATTCGAGATC 2160
Db |||||
QY 2101 ATTACAGGTGACCCCTCTCTTCTGCACTGCCCCAACTTGTCAATTCGAGATC 2160
Db |||||
QY 2161 TAATCTCCCCCTACGCTCTGCCAGGAATTTCTTTTCAGACCTCCTAGCACAGCCCGGTTG 2220
Db |||||
QY 2161 TAATCTCCCCCTACGCTCTGCCAGGAATTTCTTTTCAGACCTCCTAGCACAGCCCGGTTG 2220
Db |||||
QY 2221 CTCCTTGTGAGAGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
Db |||||
QY 2221 CTCCTTGTGAGAGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
Db |||||
QY 2281 CTCATCTTGCACCCCAACCTCTGTAATAGATTTACCGATTTTACGCTGCTGTA 2340
Db |||||
QY 2281 CTCATCTTGCACCCCAACCTCTGTAATAGATTTACCGATTTTACGCTGCTGTA 2340
Db |||||
QY 2341 GTGGGATGCTCTCTTAATGGAGAGTGTTCATTGTATTAATAGTTATTACCTGAGTAT 2400
Db |||||
QY 2341 GTGGGATGCTCTCTTAATGGAGAGTGTTCATTGTATTAATAGTTATTACCTGAGTAT 2400
Db |||||
QY 2401 GCAATAAAGATGTGGGCACTCTTTTCATGTTGGTGGGAGCAAAAAA 2456
Db |||||
QY 2401 GCAATAAAGATGTGGGCACTCTTTTCATGTTGGTGGGAGCAAAAAA 2456
Db |||||
```

RESULT 7

AEA00088
ID AEA00088 standard; cDNA; 2456 BP.

XX AEA00088;

XX 28-JUL-2005 (first entry)

XX Human TAT115 cDNA SEQ ID NO:40.

XX ss; gene; tumor-associated antigen; cytostatic; breast tumor;
XX endocrine disease; gynecology and obstetrics; neoplasm; colon tumor;
XX gastrointestinal disease; rectal tumor; endometroid carcinoma;
XX genitourinary disease; renal tumor; lung tumor; respiratory disease;
XX ovary tumor; skin tumor; liver tumor.
OS Homo sapiens.

XX	Key	Location/Qualifiers	Db	121	ACAACTGCTCAGAGTGGCGGGGATAGCTGTCCAGGTTCTCCCGGAGCTAGGAG	180
PH	CDS	254..1327	QY	181	CTCGCTGCTGCCCTCTTTCGCGCGGGAAGACAGACCAAGTTTACGCGCAACCGCTTGGC	240
FT		/*tag= a	Db	181	CTCGCTGCTGCCCTCTTTCGCGCGGGAAGACAGACCAAGTTTACGCGCAACCGCTTGGC	240
FT		/product= "TAT115"	QY	241	ACTAGGGTCCAGAAATGGGTACAAACAGTCCCTGATGGTTCCGCAATGGCTGAAATCCAA	300
XX	US2005106644-A1.		Db	241	ACTAGGGTCCAGAAATGGGTACAAACAGTCCCTGATGGTTCCGCAATGGCTGAAATCCAA	300
XX	19-MAY-2005.		QY	301	GTACTACAGACTTTGTGATAAGCTTGAAGCTTGGGGCATCGCTAGAAACGGTGGCCAC	360
XX	08-SEP-2004; 2004US-00936626.		Db	301	GTACTACAGACTTTGTGATAAGCTTGAAGCTTGGGGCATCGCTAGAAACGGTGGCCAC	360
XX	20-JUN-2001; 2001US-0299500P.		QY	361	AGCGGGGTGTGACCTCGGTGGCCCTTCACTGTCTCACTCTCCGATCCTCGTCTGCAAGGT	420
XX	29-JUN-2001; 2001US-0301880P.		Db	361	AGCGGGGTGTGACCTCGGTGGCCCTTCACTGTCTCACTCTCCGATCCTCGTCTGCAAGGT	420
XX	18-SEP-2001; 2001US-0323268P.		QY	421	GCAGGACTCCAAACAGGCGAAAAATGCTGCTACTCAGTTTCTCTCTCTCTCTCTCTCTCT	480
XX	19-JUN-2002; 2002US-00177488.		Db	421	GCAGGACTCCAAACAGGCGAAAAATGCTGCTACTCAGTTTCTCTCTCTCTCTCTCTCTCT	480
XX	26-MAR-2004; 2004US-0557116P.		QY	481	GGGCACTCTTTGGCCCTCACCTTCGCTTCACTCGGACTGGAGCGGAGCAGAGGGCCAC	540
XX	04-AUG-2004; 2004US-0598899P.		Db	481	GGGCACTCTTTGGCCCTCACCTTCGCTTCACTCGGACTGGAGCGGAGCAGAGGGCCAC	540
XX	(GETH) GENENTECH INC.		QY	541	AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTCTCTCTCTCTCTCTCTCT	600
XX	Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;		Db	541	AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTCTCTCTCTCTCTCTCTCT	600
XX	Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;		QY	601	TGTCACTGTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCCTTCTTGGTGAATCTGGG	660
XX	Sliwkowski M;		Db	601	TGTCACTGTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCCTTCTTGGTGAATCTGGG	660
XX	WPI; 2005-384304/39.		QY	661	TCTGCGCTGGGCTTCAGCTAGTCCAGGATGTTATCGTATTTGAATATATATGCTCTGAC	720
XX	P-PSDB; AEA00166.		Db	661	TCTGCGCTGGGCTTCAGCTAGTCCAGGATGTTATCGTATTTGAATATATATGCTCTGAC	720
XX	Novel isolated antibody capable of binding to tumor-associated antigenic		QY	721	CATGAATAGAGCAACAGCTCAATGTCTTTTCTGAGCTTTCGCTCTCTCGTCCGAATGAAGA	780
XX	target polypeptide, useful for treating cell proliferative disorder e.g.		Db	721	CATGAATAGAGCAACAGCTCAATGTCTTTTCTGAGCTTTCGCTCTCTCGTCCGAATGAAGA	780
XX	cancer.		QY	781	CTTTGCTCTCTGCTCAGCTAGTCTCTCTTTGATGGCGCTGACCTTCTCTATGCTCTC	840
XX	Claim 1; SEQ ID NO 40; 337pp; English.		Db	781	CTTTGCTCTCTGCTCAGCTAGTCTCTCTTTGATGGCGCTGACCTTCTCTATGCTCTC	840
XX	The invention relates to a novel isolated antibody binding to a		QY	841	CTTCACTTCTCTGCTTCTCTTCAAGGCTGGAAGACATGGGGCCCAATCTACCTCAC	900
XX	polypeptide having at least 80% sequence identity to a polypeptide having		Db	841	CTTCACTTCTCTGCTTCTCTTCAAGGCTGGAAGACATGGGGCCCAATCTACCTCAC	900
XX	any one of 76 fully defined 182-910 amino acid tumor-associated antigenic		QY	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACCTGCTCATGCTTCTGACTT	960
XX	target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the		Db	901	GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACCTGCTCATGCTTCTGACTT	960
XX	specification, a polypeptide having any one of (AEA00127-AEA00202),		QY	961	TGACGCGAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTT	1020
XX	lacking its associated signal peptide, or an extracellular domain of a		Db	961	TGACGCGAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTT	1020
XX	polypeptide having any one of (AEA00127-AEA00202). The polypeptide is		QY	1021	CTGTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1080
XX	encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).		Db	1021	CTGTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1080
XX	An antibody of the invention has cytostatic activity. The antibody is		QY	1081	TTATCCTGTGAGGATGCTTTCTGTAAACCTCAACTCTGCTGGAAGAGCTATGGTGTGA	1140
XX	useful for inhibiting growth of a cell expressing TAT188, which involves		Db	1081	TTATCCTGTGAGGATGCTTTCTGTAAACCTCAACTCTGCTGGAAGAGCTATGGTGTGA	1140
XX	contacting the cell with the antibody. The cell is a cancer cell chosen		QY	1141	GAAACAGAGCTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGACACAGGACACGCT	1200
XX	from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and		Db	1141	GAAACAGAGCTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGACACAGGACACGCT	1200
XX	liver cell. The cancer cell is a mammalian cell, preferably a human cell.		QY	1201	CTATGCGCCCTTATTTCCACACATTTTTCAGCTGAGAACACAGCTTCCCGGAGGAAATTTCTC	1260
XX	The antibody is also useful for detecting the level of TAT188		Db			
XX	polypeptide expressed in a test cell relative to a control cell, and for		QY			
XX	detecting the level of TAT188 polypeptide or a polypeptide having at		Db			
XX	least 80 % sequence identity to the TAT188 polypeptide sequence in a test		QY			
XX	cell relative to a control cell. The antibody is useful for treating a		Db			
XX	preventing a cell proliferative disorder associated with increased		QY			
XX	expression or activity of a polypeptide having at least 80 % identity to		Db			
XX	a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.		QY			
XX	The method of the invention is useful for inhibiting the growth of a		Db			
XX	cancer cell. The present sequence encodes a polypeptide of the invention.		QY			
XX	Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;		QY			
XX	Query Match 100.0%; Score 2456; DB 14; Length 2456;		QY			
XX	Best Local Similarity 100.0%; Pred. No. 0;		QY			
XX	Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		QY			
QY	1 ATAAACAGCATGAAGTCCGCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC	60	QY	1	ATAACAGCATGAAGTCCGCTGGAACCTGGAATAGGCGTGTCTCTCTCCCTCGACCCCTCCCC	60
Db	1 ATAAACAGCATGAAGTCCGCTGGAACCTGGAATAGGCGTGTCTCTCTCCCTCGACCCCTCCCC	60	QY	61	TCTTTGCTCTGCTCACCCTTCGCTCGTTCCTCCCTCCGCGAGGCGCGCTTTATA	120
QY	61 TCTTTGCTCTGCTCACCCTTCGCTCGTTCCTCCCTCCGCGAGGCGCGCTTTATA	120	Db	61	TCTTTGCTCTGCTCACCCTTCGCTCGTTCCTCCCTCCGCGAGGCGCGCTTTATA	120
Db	61 TCTTTGCTCTGCTCACCCTTCGCTCGTTCCTCCCTCCGCGAGGCGCGCTTTATA	120	QY	121	ACAACTGCTCAGAGTGGGGGGGATAGCTGTCCAAGGTTCTCCCGGAGCTAGGAG	180

Db 1201 CTATGCCCTTATTCACACATTTTCAGCTGCAGAACCGCTCCCCAAAAGGAATTC 1260
Qy 1261 CATCCACGGCCCGACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAGAAAGAGG 1320
Db 1261 CATCCACGGCCCGACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAGAAAGAGG 1320
Qy 1321 CAGCTAACTCTGTCTGAAGAGTGGGACAAATGACGCGGGCGGAGATCTAGCGGAGC 1380
Db 1321 CAGCTAACTCTGTCTGAAGAGTGGGACAAATGACGCGGGCGGAGATCTAGCGGAGC 1380
Qy 1381 TCAAGGGATGTGGGCGAATCTTGAGTCTTCTGAGAAAACCTCTACAAGACACTACGGGA 1440
Db 1381 TCAAGGGATGTGGGCGAATCTTGAGTCTTCTGAGAAAACCTCTACAAGACACTACGGGA 1440
Qy 1441 ACAGTTTGCTCCCTCCAGCTCAACCAAAATCTTCCATGCTGGGCTGATGGGCT 1500
Db 1441 ACAGTTTGCTCCCTCCAGCTCAACCAAAATCTTCCATGCTGGGCTGATGGGCT 1500
Qy 1501 AGTAAGACTCCAGTCTCTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
Db 1501 AGTAAGACTCCAGTCTCTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
Qy 1561 ATACTCTTTTAAAGTGGAGTCTCAGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
Db 1561 ATACTCTTTTAAAGTGGAGTCTCAGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
Qy 1621 GTTTTTTGAACAGGATCTGCTCTCACCCAGGCTTGAGTCAGTGGGATCACAG 1680
Db 1621 GTTTTTTGAACAGGATCTGCTCTCACCCAGGCTTGAGTCAGTGGGATCACAG 1680
Qy 1681 CCAGTGCAGCCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db 1681 CCAGTGCAGCCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Qy 1741 CTGGATGACAGCGTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTGTTATTT 1800
Db 1741 CTGGATGACAGCGTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTGTTATTT 1800
Qy 1801 CCATGACTAAAGGTCTGCTCATCTGAGTCAAGTGGCTCAGAGCTCTAGGGCCTG 1860
Db 1801 CCATGACTAAAGGTCTGCTCATCTGAGTCAAGTGGCTCAGAGCTCTAGGGCCTG 1860
Qy 1861 CTCCTCTAACTCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGAGCTCATATCT 1920
Db 1861 CTCCTCTAACTCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGAGCTCATATCT 1920
Qy 1921 GAGCAAAATAGCAAAAGGCTCTCTCAGCCCACTGGGCTGAATCTACACTGGAAGCCAAC 1980
Db 1921 GAGCAAAATAGCAAAAGGCTCTCTCAGCCCACTGGGCTGAATCTACACTGGAAGCCAAC 1980
Qy 1981 TTGCTGGCACCCGCTCCCAACCTTCTTGCTGGTAGGAGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCGCTCCCAACCTTCTTGCTGGTAGGAGGCTAAAGATCACCC 2040
Qy 2041 TAAATTTACTCATCTCTAGTCTGCCTCACATTTGGGCTCAGAGCTCCCGAGCACA 2100
Db 2041 TAAATTTACTCATCTCTAGTCTGCCTCACATTTGGGCTCAGAGCTCCCGAGCACA 2100
Qy 2101 ATTACAGGTACCCCTCTCTTCTTGCACTGTCCCAAACTGCTGTCAATCCGAGATC 2160
Db 2101 ATTACAGGTACCCCTCTCTTCTTGCACTGTCCCAAACTGCTGTCAATCCGAGATC 2160
Qy 2161 TAATCTCCCTACGCTCTGGAGGAATTTTTCAGACCTCAGTACAGACGCGGGTGG 2220
Db 2161 TAATCTCCCTACGCTCTGGAGGAATTTTTCAGACCTCAGTACAGACGCGGGTGG 2220
Qy 2221 CTCCTTGTAGGAGATTTGTAGATCATCTCACTTCAATTCCTGGGCTGATCTCT 2280
Db 2221 CTCCTTGTAGGAGATTTGTAGATCATCTCACTTCAATTCCTGGGCTGATCTCT 2280
Qy 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTTACCGATTTTACGGCTGCAATCTGTAA 2340
Db 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTTTACCGATTTTACGGCTGCAATCTGTAA 2340

Qy 2341 GTGGCATGCTCTCTAATGGAGGAGTGTTCATTGTATATAGTTATTCCACCTGAGTAT 2400
Db 2341 GTGGCATGCTCTCTAATGGAGGAGTGTTCATTGTATATAGTTATTCCACCTGAGTAT 2400
Qy 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456

RESULT 8
AEA00608
ID AEA00608 standard; cDNA; 2456 BP.
XX AEA00608;
AC AEA00608;
XX DT 28-JUL-2005 (first entry)
XX Human TAT115 cDNA sequence SeqID40.
XX antibody identification; tumor-associated antigen; cytostatic;
KW RNA interference; gene therapy; cell death; cancer; breast tumor;
KW colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;
KW skin tumor; liver tumor; gene; ss; TAT115.
XX OS Homo sapiens.
XX US2005107595-A1.
XX PD 19-MAY-2005.
XX PF 10-SEP-2004; 2004US-00938061.
XX PR 20-JUN-2001; 2001US-0299500P.
PR 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323268P.
PR 19-JUN-2002; 2002US-00177488.
PR 26-MAR-2004; 2004US-0557116P.
PR 04-AUG-2004; 2004US-0598899P.
XX (GETH) GENENTECH INC.
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sakanaka C, Chuntharapai A, Reed CJ;
XX WPI; 2005-371577/38.
DR P-PSDB; AEA00686.
PT Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds
PT to tumor-associated antigenic target polypeptide, useful for diagnosing
XX or treating cancer.
PS Claim 1; SEQ ID NO 40; 96pp; English.
XX This invention relates to a novel isolated antibody, for example anti-E16
or anti-TAT112 antibody, that binds to a tumor-associated antigenic
target polypeptide (TAT) and that lacks an associated signal peptide
sequence. The invention may be useful for the development of compounds
with a cytostatic activity acting as antagonists of the TAT188
polypeptide or RNA interference whilst the disclosed sequences may be
useful for gene therapy. The invention is useful for inducing the death
of a cell (such as a cancer cell chosen from breast, colon, rectum,
endometrium, kidney, lung, ovary, skin and liver) to which it binds,
inhibiting proliferation or promoting cell death of a cell expressing
TAT188. In addition, the invention may be useful for detecting the level
of TAT188 polypeptide in a test cell relative to a control cell, or
treating or preventing a cell proliferative disorder associated with
increased expression of TAT188. The novel antibody of the invention is
useful for inhibiting the growth of a cancer cell and may be useful for
diagnosing or treating cancer. The present sequence is that of the human
TAT115 cDNA which encodes a protein against which an antibody of the
invention may be targeted.

XX	Sequence	2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;			
SQ	Query Match	100.0%; Score 2456; DB 14; Length 2456;			
	Best Local Similarity	100.0%; Pred. No. 0;			
	Matches 2456; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATAACAGCATGAAGTGCCTCGAACTGGAATAGGCGTGTCTCTCCCTCGGCGAGGGCGCCCTTATA	120		
DB	1	ATAACAGCATGAAGTGCCTCGAACTGGAATAGGCGTGTCTCTCCCTCGAACCCTCCCC	60		
QY	61	TCCTTGTCCCTCTGCTCACCCCTCGTCTTCCCTCCCTCCGCGAGGGCGCCCTTATA	120		
DB	61	TCCTTGTCCCTCTGCTCACCCCTCGTCTTCCCTCCCTCCGCGAGGGCGCCCTTATA	120		
QY	121	ACAACTGCTCAGAGTGCAGAGGGCGGATAGCTGTCCAAGGTCTCCGCCAGCACTGAGGAG	180		
DB	121	ACAACTGCTCAGAGTGCAGAGGGCGGATAGCTGTCCNAGGTCTCCGCCAGCACTGAGGAG	180		
QY	181	CTGGCTGTGCTTCTGCGCGCGGGAAGAGCAACAAAGTTACGGCCAAACGGCTTGGC	240		
DB	181	CTCGCTGTGCTTCTGCGCGCGGGAAGAGCAACAAAGTTACGGCCAAACGGCTTGGC	240		
QY	241	ACTAGGCTCCAGATGGCTACACAGTCCCTGATGGTTGCGCAATGGCTGAATCCAA	300		
DB	241	ACTAGGCTCCAGATGGCTACACAGTCCCTGATGGTTGCGCAATGGCTGAATCCAA	300		
QY	301	GTACTACAGACTTTGTGATAAGCTGAAGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360		
DB	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360		
QY	361	AGCGGGGTGTGACCTCGGTGGCTTATGCTCACTCTCCGATCTCGTCTGCAAGGT	420		
DB	361	AGCGGGGTGTGACCTCGGTGGCTTATGCTCACTCTCCGATCTCGTCTGCAAGGT	420		
QY	421	GCAGACTCCACAGCGGAAATGCTGCTACTCAGTTTCTCTTCTCTCGGTGTGT	480		
DB	421	GCAGACTCCACAGCGGAAATGCTGCTACTCAGTTTCTCTTCTCTCGGTGTGT	480		
QY	481	GGGCATCTTTGGCTCACCTTCGCCTTCATCATCGGACTGGAGCGGAGCACAGGGCCAC	540		
DB	481	GGGCATCTTTGGCTCACCTTCGCCTTCATCATCGGACTGGAGCGGAGCACAGGGCCAC	540		
QY	541	AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTGCTCATGC	600		
DB	541	AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTGCTCATGC	600		
QY	601	TGTCAGTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCCTGTTGGTGAATCTGGG	660		
DB	601	TGTCAGTCTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCCTGTTGGTGAATCTGGG	660		
QY	661	TCCTGGCCGTGGGCTTCAGGCTAGTCAGGATGTATCGCTATTGAATATATTGTCTCGAC	720		
DB	661	TCCTGGCCGTGGGCTTCAGGCTAGTCAGGATGTATCGCTATTGAATATATTGTCTCGAC	720		
QY	721	CATGAATAGGACCAACAGTCAATGTCTTTTCTGAGCTTTCGGTCTCTCGTCAATGAAGA	780		
DB	721	CATGAATAGGACCAACAGTCAATGTCTTTTCTGAGCTTTCGGTCTCTCGTCAATGAAGA	780		
QY	781	CTTTGTCCTCTGCTCACCTACGTCCTCTTCTTGATGGCGCTGACCTTCTCATGTCCTC	840		
DB	781	CTTTGTCCTCTGCTCACCTACGTCCTCTTCTTGATGGCGCTGACCTTCTCATGTCCTC	840		
QY	841	CTTTCACCTTCTGTGGTTCCTTTCACGGGCTGGAGAGACATGGGGCCCAATCTACCTCAC	900		
DB	841	CTTTCACCTTCTGTGGTTCCTTTCACGGGCTGGAGAGACATGGGGCCCAATCTACCTCAC	900		
QY	901	GATGCTCTCTCGATTGGCCATCTGGGTGGCTGGATCACCTGCTCATGCTCTCTGACTTT	960		
DB	901	GATGCTCTCTCGATTGGCCATCTGGGTGGCTGGATCACCTGCTCATGCTCTCTGACTTT	960		
QY	961	TCACCGCAGGTGGGATGACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGT	1020		
DB	961	TCACCGCAGGTGGGATGACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGT	1020		

961	TCACCGCAGGTGGGATGACCATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGT	1020
1021	CCTGTTGGCTTATGTTAGTCCGAGTTTTGGCTGTCTCAAAAGCAACGAAACCCCATGGA	1080
1021	CCTGTTGGCTTATGTTAGTCCGAGTTTTGGCTGTCTCAAAAGCAACGAAACCCCATGGA	1080
1081	TTATCCTGTTGAGGATGCTTCTGTAAACCTCAACTCGTGTGAAGAGAGCTATGGTGTGA	1140
1081	TTATCCTGTTGAGGATGCTTCTGTAAACCTCAACTCGTGTGAAGAGAGCTATGGTGTGA	1140
1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGTTTTTGAAGAGACAGGGGACACGCT	1200
1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGTTTTTGAAGAGACAGGGGACACGCT	1200
1201	CTATGCCCTTATTCACACATTTTTCAGCTGCAGAACCAAGCTTCCCAAAAGGAAATTC	1260
1201	CTATGCCCTTATTCACACATTTTTCAGCTGCAGAACCAAGCTTCCCAAAAGGAAATTC	1260
1261	CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG	1320
1261	CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGG	1320
1321	CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGACGCGGGCGGACAGATCTAGCGGAGC	1380
1321	CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGACGCGGGCGGACAGATCTAGCGGAGC	1380
1381	TCAAAAGGAGTGGGGGAAATCTTGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGA	1440
1381	TCAAAAGGAGTGGGGGAAATCTTGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGA	1440
1441	ACAGTTTGGCTCCCTCCAGCCCTCAACCAATTTTCCATGCTGGGGCTGATGGGCT	1500
1441	ACAGTTTGGCTCCCTCCAGCCCTCAACCAATTTTCCATGCTGGGGCTGATGGGCT	1500
1501	AGTAAGACTCCAGTTCTTAGAGGCGCTGAGTATTTTTTTTTTTTTTGTCTCATCTTTGG	1560
1501	AGTAAGACTCCAGTTCTTAGAGGCGCTGAGTATTTTTTTTTTTTTTGTCTCATCTTTGG	1560
1561	ATACTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTATAGACCTTACTCTTTTGT	1620
1561	ATACTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTATAGACCTTACTCTTTTGT	1620
1621	GTTTTTTGAACAGAGATCTTGTCTGTACCCAGGCTTGAAGTGGATGGATCACAG	1680
1621	GTTTTTTGAACAGAGATCTTGTCTGTACCCAGGCTTGAAGTGGATGGATCACAG	1680
1681	CCAGTGCAGCTCGACACCTGTGTCAAGCAATCTCCCATCTCCCAAGTG	1740
1681	CCAGTGCAGCTCGACACCTGTGTCAAGCAATCTCCCATCTCCCAAGTG	1740
1741	CTGGGATGACAGGCGTGAGCCACAGCTCCAGCCCTAGGCCCTTAAATCTTGTCTTATTT	1800
1741	CTGGGATGACAGGCGTGAGCCACAGCTCCAGCCCTAGGCCCTTAAATCTTGTCTTATTT	1800
1801	CATGGAATAAGGCTTGTGATCTGAGCTCAGCTGGCTCACACAGCTCTAGGGCCCTG	1860
1801	CATGGAATAAGGCTTGTGATCTGAGCTCAGCTGGCTCACACAGCTCTAGGGCCCTG	1860
1861	CTCCTCTAACTCACAAGTGGGTTTTGTGAGGCTCTGTGGCCACAGACAGCTGCTATCT	1920
1861	CTCCTCTAACTCACAAGTGGGTTTTGTGAGGCTCTGTGGCCACAGACAGCTGCTATCT	1920
1921	GAGCAAAAATAGCAAAAGCCTCTCAGCCCACTGGCTGAAATCTACACTGGAAGCAAC	1980
1921	GAGCAAAAATAGCAAAAGCCTCTCAGCCCACTGGCTGAAATCTACACTGGAAGCAAC	1980
1981	TTGCTGGCACCCCGCTCCCAACCTTCTTCCCTGGGTAGGAGGCTAAAGATCACCC	2040
1981	TTGCTGGCACCCCGCTCCCAACCTTCTTCCCTGGGTAGGAGGCTAAAGATCACCC	2040
2041	TAAATTTACTCATCTCTTAGTGCTGCTCACAATGGGCTCAGCAGCTCCCGCAGACCA	2100
2041	TAAATTTACTCATCTCTTAGTGCTGCTCACAATGGGCTCAGCAGCTCCCGCAGACCA	2100


```
QY 2101 ATTACAGGTACCCCTCTCTTCTTGACATGTCCTCCCAAACTTGCTGTCAATTCGGAGTC 2160
Db |||||||
QY 2101 ATTACAGGTACCCCTCTCTTCTTGACATGTCCTCCCAAACTTGCTGTCAATTCGGAGTC 2160
Db |||||||
QY 2161 TAATCTCCCTACGCTCTGCCAGGAATTTCTTTTCAGACCTCACTAGCACAAAGCCCGTTG 2220
Db |||||||
QY 2161 TAATCTCCCTACGCTCTGCCAGGAATTTCTTTTCAGACCTCACTAGCACAAAGCCCGTTG 2220
Db |||||||
QY 2221 CTCTTGTTCAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATTTCTGGGGCTGATCTTCT 2280
Db |||||||
QY 2221 CTCTTGTTCAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATTTCTGGGGCTGATCTTCT 2280
Db |||||||
QY 2281 CTCACTTGTGACCCCAACCTCTGTAATAGATTTTACCGCATTTTACGCTCATTTCTGTAA 2340
Db |||||||
QY 2281 CTCACTTGTGACCCCAACCTCTGTAATAGATTTTACCGCATTTTACGCTCATTTCTGTAA 2340
Db |||||||
QY 2341 GTGGGATGTCTCTTAATGAGGAGTGTTTCATTGTATAATAAGTTATTCACCTGAGTAT 2400
Db |||||||
QY 2341 GTGGGATGTCTCTTAATGAGGAGTGTTTCATTGTATAATAAGTTATTCACCTGAGTAT 2400
Db |||||||
QY 2401 GCAATAAGATGTGGTGGCCACTCTTTCATGTTGGTGGCAGCAAAAAAAAA 2456
Db |||||||
QY 2401 GCAATAAGATGTGGTGGCCACTCTTTCATGTTGGTGGCAGCAAAAAAAAA 2456
Db |||||||

RESULT 9
AED47480
ID AED47480 standard; DNA; 2456 BP.
XX
AC AED47480;
XX
DT 15-DEC-2005 (first entry)
XX
DE Retinoic acid-inducible G-protein coupled receptor 3 gene, SEQ ID 2.
XX
KW RNA interference; gene silencing;
KW retinoic acid-inducible G-protein coupled receptor 3; breast tumor;
KW cytostatic; cell growth; gene; ds.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FH 254..1327
FH CDS
FH /*tag= a
FH /product= "Retinoic acid-inducible G-protein coupled
FH receptor 3"
FH /*tag= b
FH /label= Transmembrane domain
FH 452..514
FH /*tag= c
FH /label= Transmembrane domain
FH 497..1057
FH /*tag= d
FH /label= 7 transmembrane receptor
FH 548..601
FH /*tag= e
FH /label= Transmembrane domain
FH 656..718
FH /*tag= f
FH /label= Transmembrane domain
FH 776..832
FH /*tag= g
FH /label= Transmembrane domain
FH 887..946
FH /*tag= h
FH /label= Transmembrane domain
FH 2403..2408
FH /*tag= i
FH polyA_site
FH 2442..2456
FH /*tag= j
XX
```

```
PN JP2005287456-A.
XX
PD 20-OCT-2005.
XX
PF 02-APR-2004; 2004JP-00110232.
XX
PR 02-APR-2004; 2004JP-00110232.
XX
PA (NIPK ) NIPPON KAYAKU KK.
PA (UYNI-) UNIV NIPPON IKA.
XX
PI Nagahata T, Emi M;
XX
DR WPI: 2005-737982/76.
DR P-PSDB; AED47479.
XX
PT Substance that suppresses expression of retinoic acid-inducible G-protein
PT coupled receptor-3 RA13 gene or nucleic acid, useful for suppressing cell
PT growth, and as pharmaceutical or therapeutic agent of breast cancer.
XX
PS Claim 1; SEQ ID NO 2; 22pp; Japanese.
XX
CC The invention relates to a novel substance which suppresses the
CC expression of a retinoic acid-inducible G-protein coupled receptor (RA1)3
CC gene. The RA13 gene suppressing agent is useful for suppressing cell
CC growth, and as a pharmaceutical or therapeutic agent of breast cancer.
CC and is also useful in identifying a cell growth inhibitory substance. The
CC RA13 gene suppressing agent enables the identification of a cell growth
CC inhibitory substance. This polynucleotide sequence represents the
CC retinoic acid-inducible G-protein coupled receptor 3 gene of the
CC invention.
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 2456; DB 14; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAACAGCATGAAGTCCCTGGAACCTGGAATAGGCGGTCTCTCCCTCGACCCCTCCCC 60
Db |||||||
QY 61 TCCTTGTCCCTCTGCTCACCCCTCGCTTCCCTCCCGGAGGGCGGCTTTTATA 120
Db |||||||
QY 61 TCCTTGTCCCTCTGCTCACCCCTCGCTTCCCTCCCGGAGGGCGGCTTTTATA 120
Db |||||||
QY 121 ACAACTGTCTCAGAGTCCGAGGGCGGATAGCTGTCCAAGGTCTCCCGCAGCACTGAGGAG 180
Db |||||||
QY 181 CTGCGCTGTCCCTCTTTCGCGCGGGAAGCAGCACAAGTTTACGGCCAAAGCCCTTGGC 240
Db |||||||
QY 241 ACTAGGTCACAGAAATGGCTACCAACAGTCCCTCATGTGTTCGCGCAATGGCTTGAATCCAA 300
Db |||||||
QY 301 GTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATCGTCTAGAAAACGTTGCCAC 360
Db |||||||
QY 361 AGCCGGGTGTGACCTCGGTGGCCCTTCATGCTCACTCTCCGATCTCGTCTGCAAGGT 420
Db |||||||
QY 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCTAGTTTCTCTTCTCTCTGCTGTT 480
Db |||||||
QY 481 GGGCATCTTTGGCCTCACCTTCGCCCTTCATTCGGAAGTGGAGCGGAGCAGAGGGCCAC 540
Db |||||||
QY 481 GGGCATCTTTGGCCTCACCTTCGCCCTTCATTCGGAAGTGGAGCGGAGCAGAGGGCCAC 540
Db |||||||
```

541 QY ACCTCTTCTTCCCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTCTGCTCTGCTCATGC 600
541 Db AGCTTCTTCTTCTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTCTGCTCTGCTCATGC 600
601 QY TGTCACTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGAATTCGGG 660
601 Db TGTCACTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGAATTCGGG 660
661 QY TCTGGCCGTGGGCTTCAGCCTAGTCAGGATGTTATCGCTATTGAATATATTGTCCTGAC 720
661 Db TCTGGCCGTGGGCTTCAGCCTAGTCAGGATGTTATCGCTATTGAATATATTGTCCTGAC 720
721 QY CATGAATAGGACCAACGTCATATGTCCTTTCTGAGCTTTCCGCTCTCTGTCGCAATGAAGA 780
721 Db CATGAATAGGACCAACGTCATATGTCCTTTCTGAGCTTTCCGCTCTCTGTCGCAATGAAGA 780
781 QY CTTTGTCTCTGCTCCTACCTAGTCCTCTCTTCTGATGGGCTGACCTTCTCATGTCCTC 840
781 Db CTTTGTCTCTGCTCCTACCTAGTCCTCTCTTCTGATGGGCTGACCTTCTCATGTCCTC 840
841 QY CTTTACCTTCTGTGGTTCCTTCAAGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900
841 Db CTTTACCTTCTGTGGTTCCTTCAAGGCTGGAAGAGACATGGGGCCCAATCTACCTCAC 900
901 QY GATGCTCTCTCATTTGCCATCTGGGTGCTGGATCAACCTGCTCATGCTTCTGACTT 960
901 Db GATGCTCTCTCATTTGCCATCTGGGTGCTGGATCAACCTGCTCATGCTTCTGACTT 960
961 QY TGACCCAGGTGGGATGACCAATCTCAGCTCCGCTTGGCTGCGCAATGGCTGGGTGTT 1020
961 Db TGACCCAGGTGGGATGACCAATCTCAGCTCCGCTTGGCTGCGCAATGGCTGGGTGTT 1020
1021 QY CTTGTTGGCTTATGTTAGTCCGAGTTTGGGTGCTCACAAGCAACGAAACCCCATGGA 1080
1021 Db CTTGTTGGCTTATGTTAGTCCGAGTTTGGGTGCTCACAAGCAACGAAACCCCATGGA 1080
1081 QY TTATCTCTGTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
1081 Db TTATCTCTGTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
1141 QY GAACAGAGCTACTCTCAAGAGAAATCACTCAAGCTTTGAAGACAGGGGACACGCT 1200
1141 Db GAACAGAGCTACTCTCAAGAGAAATCACTCAAGCTTTGAAGACAGGGGACACGCT 1200
1201 QY CTATGCCCTCTATTTCACACATTTTCAGCTGCAGAACCCGCTCCCAAGGAATTCCT 1260
1201 Db CTATGCCCTCTATTTCACACATTTTCAGCTGCAGAACCCGCTCCCAAGGAATTCCT 1260
1261 QY CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAGAAAGAGG 1320
1261 Db CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAGAAAGAGG 1320
1321 QY CAGCTAACTCTGCTCAAGAGTGGGACAAATGSCAGCCGGGCGAGATCTAGCGGAGC 1380
1321 Db CAGCTAACTCTGCTCAAGAGTGGGACAAATGSCAGCCGGGCGAGATCTAGCGGAGC 1380
1381 QY TCAAGGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
1381 Db TCAAGGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
1441 QY ACAGTTTGGCTCCCTCCAGCTCAACCAATTCATGCTGGGCTGATGGGCT 1500
1441 Db ACAGTTTGGCTCCCTCCAGCTCAACCAATTCATGCTGGGCTGATGGGCT 1500
1501 QY AGTAAGACTCCAGTCTTAGAGCGCTGTAGTATTTTCTTCTCATCCCTTGG 1560
1501 Db AGTAAGACTCCAGTCTTAGAGCGCTGTAGTATTTTCTTCTCATCCCTTGG 1560
1561 QY ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
1561 Db ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620

1621 QY GTTTTGTGAAACAGGATCTTGCTCTGTCAACCAGGCTTGAGTCCAGTGGATCACAG 1680
1621 Db GTTTTGTGAAACAGGATCTTGCTCTGTCAACCAGGCTTGAGTGGATCACAG 1680
1681 QY CCAGTGCAGCCTCGACACCTGTGCTCAAGCAATCCTCCCATCTCCATCTCCCAAGTG 1740
1681 Db CCAGTGCAGCCTCGACACCTGTGCTCAAGCAATCCTCCCATCTCCATCTCCCAAGTG 1740
1741 QY CTGGGATGACAGCGGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTGTGTTATTT 1800
1741 Db CTGGGATGACAGCGGTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTGTGTTATTT 1800
1801 QY CCATGACTAAAGGCTTGCTCATCTGAGCTCAGCTGAGCTCAGAGCTCTAGGGCCCTG 1860
1801 Db CCATGACTAAAGGCTTGCTCATCTGAGCTCAGCTGAGCTCAGAGCTCTAGGGCCCTG 1860
1861 QY CTCCTCTAACTCACAGTGGGTTTGTGAGGCTCTGTGGGCCAGAGCAGACCTGCTATCT 1920
1861 Db CTCCTCTAACTCACAGTGGGTTTGTGAGGCTCTGTGGGCCAGAGCAGACCTGCTATCT 1920
1921 QY GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGCGCTGAATCTACACTGGAAGCAAC 1980
1921 Db GAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGCGCTGAATCTACACTGGAAGCAAC 1980
1981 QY TTGCTGGCACCCCGCTCCCAACCTTCTTGGCTGGGTAGGAGGCTAAGATCACCC 2040
1981 Db TTGCTGGCACCCCGCTCCCAACCTTCTTGGCTGGGTAGGAGGCTAAGATCACCC 2040
2041 QY TAAATTTACTCATCTCTTAGTGTGCTCACAATGGGCTCAGCAGCTCCCCAGCACCA 2100
2041 Db TAAATTTACTCATCTCTTAGTGTGCTCACAATGGGCTCAGCAGCTCCCCAGCACCA 2100
2101 QY ATTACAGTCAACCCCTCTCTTCTTGCACTGCCCCAAACTTGTCTCAATTCGAGATC 2160
2101 Db ATTACAGTCAACCCCTCTCTTCTTGCACTGCCCCAAACTTGTCTCAATTCGAGATC 2160
2161 QY TAACTCTCCCTACGCTCTGCCAGGAATCTTTCAGACCTCACTAGCAAGCCCGGTTG 2220
2161 Db TAACTCTCCCTACGCTCTGCCAGGAATCTTTCAGACCTCACTAGCAAGCCCGGTTG 2220
2221 QY CTCCTTGTGAGGAAATTTGTAGATCATTTCTCACTTCAAATTCCTGGGCTGATCTCT 2280
2221 Db CTCCTTGTGAGGAAATTTGTAGATCATTTCTCACTTCAAATTCCTGGGCTGATCTCT 2280
2281 QY CTCATCTTGCAACCCCAACCTCTGTAAATAGATTTACCGATTTACGGCTGCAATCTGTA 2340
2281 Db CTCATCTTGCAACCCCAACCTCTGTAAATAGATTTACCGATTTACGGCTGCAATCTGTA 2340
2341 QY GTGGGATGCTCTCCTTAATGGAGGAGTTCATTTGATATAAGTTATTACCTGAGTAT 2400
2341 Db GTGGGATGCTCTCCTTAATGGAGGAGTTCATTTGATATAAGTTATTACCTGAGTAT 2400
2401 QY GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456
2401 Db GCAATAAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCAAAAAA 2456

RESULT 10

ADI28525
ID ADI28525 standard; cDNA; 2456 BP.

XX AC ADI28525;

XX DT

XX DT 22-APR-2004 (first entry)

XX Human GPCR retinoic acid induced 3 (RAI-3) cDNA.

DE Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
XX antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
KW chromosome 12p13-p12.3; ss.
XX Homo sapiens.

OS


```

XX Key
FH variation
FT Location/Qualifiers
FT replace(112,r)
FT /*tag= a
FT /label= RAI-3-s1
FT /note= "located in 5' untranslated region"
FT /standard_name= "Single nucleotide polymorphism"
FT CDS
FT 254..1327
FT /*tag= b
FT /product= "Human RAI3"
FT replace(364,y)
FT /*tag= c
FT /label= RAI-3-s2
FT /note= "exon 1, silent (Ala/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation
FT replace(511,y)
FT /*tag= d
FT /label= RAI-3-s3
FT /note= "exon 2, silent (Ile/Ile)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation
FT replace(523,y)
FT /*tag= e
FT /label= RAI-3-s4
FT /note= "exon 2, silent (Asp/Asp)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation
FT replace(605,r)
FT /*tag= f
FT /label= RAI-3-s6
FT /note= "exon 2, missense (Ser/Gly)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation
FT replace(797,r)
FT /*tag= g
FT /label= RAI-3-s5
FT /note= "exon 2, missense (Thr/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation
FT replace(1111,r)
FT /*tag= h
FT /label= RAI-3-s8
FT /note= "silent (Pro/Pro)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation
FT replace(1173,r)
FT /*tag= i
FT /label= RAI-3-s9
FT /note= "missense (Gln/Arg)"
FT /standard_name= "Single nucleotide polymorphism"
FT W02004001060-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019255.
XX
XX 20-JUN-2002; 2002US-0390850P.
XX 29-AUG-2002; 2002US-0407006P.
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Whitney GS, Opitck G, Garulacan L, Ramanathan CS, McKinnon M;
XX Bennett KL, Barber UE, Cacace A, Tsuchihashi Z;
XX WPI; 2004-090973/09.
XX P-PSDB; ADI28526.
XX
XX New nucleic acid molecule encoding a human G-protein coupled receptor,
XX RAI3, useful for preventing, treating or ameliorating chronic obstructive
XX pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
XX of COPD.
XX
XX Claim 21; SEQ ID NO 18; 301pp; English.
XX
XX The present sequence is that of cDNA encoding a human G-protein coupled
XX receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
XX Proteomics methods were used to isolate cigarette smoke-inducible

```

tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was identified as being tyrosine phosphorylated and/or as being associated/complexed with tyrosine phosphorylated proteins only in those cells that had been exposed to cigarette smoke. Since RAI-3 is primarily expressed in lung tissue, and since cigarette smoke is a major causative factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a novel cellular target for identifying modulators, e.g. agonists or antagonists, useful for the treatment and/or prevention of COPD and related disorders such as emphysema and chronic bronchitis. RAI-3 modulators, e.g. agonists and antagonists, especially antisense compounds, can be used to treat COPD and other disorders and diseases associated with regulation of NF-kB and/or its associated or interacting signaling molecules. Single nucleotide polymorphisms (SNPs) detected in the RAI-3 gene are useful for determining COPD association in individuals. RAI-3 nucleic acid molecules and polypeptides are useful for preventing, treating or ameliorating disorders related to aberrant GPCR signaling or cell cycle regulation, pulmonary disorders, inflammatory lung disorders, COPD, the underlying symptoms of COPD, COPD-related disorders or conditions, autoimmune disorders, disorders related to hyperimmune activity, inflammatory conditions, disorders related to aberrant acute phase responses, hypercongenital conditions, birth defects, necrotic lesions, wounds, organ transplant rejection, renal diseases, ischaemia-reperfusion injury, heart disorders, disorders related to aberrant signal transduction, proliferation disorders, cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis, ulcerative colitis, cerebral infarct, myocardial infarct, diabetic nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis, rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma, pulmonary small cell undifferentiated carcinoma, carcinoma of the breast, colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders associated with aberrant cell adhesion, I-CAM function and/or regulation, E-selectin function and/or regulation, or aberrant NF-kB function and/or regulation (all claimed).

XX
 SQ Sequence 2456 BP; 516 A; 719 C; 570 G; 643 T; 0 U; 8 Other;

	Query Match	99.7%	Score 2448	DB 12	Length 2456
	Best Local Similarity	99.7%	Pred. No. 0		
	Matches 2448	Conservative 0	Mismatches 8	Indels 0	Gaps 0
QY	1	ATAACAGCATGAAGTCGCGTGGAACTGGAAATAGGCGGTGCTCTCCCTCGACCCCTCCCC	60		
DB	1	ATAACAGCATGAAGTCGCGTGGAACTGGAAATAGGCGGTGCTCTCCCTCGACCCCTCCCC	60		
QY	61	TCCTTGTCCCTCTGCTCAGCCCTCGCTCGCTTCCTCCCTCCGCGAGGCGCCCTTTATA	120		
DB	61	TCCTTGTCCCTCTGCTCAGCCCTCGCTCGCTTCCTCCCTCCGCGAGGCGCCCTTTATA	120		
QY	121	ACAACTGCTCAGAGTCGAGGGGGGATAGCTGCCAAGGTCTCCCAAGGTCTCCCCAGCACTGAGGAG	180		
DB	121	ACAACTGCTCAGAGTCGAGGGGGGATAGCTGCCAAGGTCTCCCAAGGTCTCCCCAGCACTGAGGAG	180		
QY	181	CTGCGCTGTGCCCTCTTGGCGCGGGGAAGCAGCACCAAGTTCCAGGCTTCCCGCCAAACGCTTGGC	240		
DB	181	CTGCGCTGTGCCCTCTTGGCGCGGGGAAGCAGCACCAAGTTCCAGGCTTCCCGCCAAACGCTTGGC	240		
QY	241	ACTAGGGTCAGAAATCGCTACACAGTCCTGATGTTGTCGCGCAATGGCCCTGAAATCCAA	300		
DB	241	ACTAGGGTCAGAAATCGCTACACAGTCCTGATGTTGTCGCGCAATGGCCCTGAAATCCAA	300		
QY	301	GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360		
DB	301	GTACTACAGACTTTGTGATAAGGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360		
QY	361	AGCCGGGGTTGTGACCTCGGTGGCCCTTCACTGCTCCTCCGATCCTCGTCTGCAAGGT	420		
DB	361	AGCCGGGGTTGTGACCTCGGTGGCCCTTCACTGCTCCTCCGATCCTCGTCTGCAAGGT	420		
QY	421	GCAGGACTCCACAGCGGAAAAATGCTGCTACTCAGTTTCTCTTCTCTCTGGGTGTT	480		
DB	421	GCAGGACTCCACAGCGGAAAAATGCTGCTACTCAGTTTCTCTTCTCTCTGGGTGTT	480		
QY	481	GGGCATCTTTGGCCTCACCTTCGCTTCATCATCGGACTGACGGGAGCAGAGGGCCAC	540		

PN	US2003099974-A1.	Db	889	GGGCACTCTTTGGGCTCACCTTCGCTTCATCATCGGACTGAGCGGAGACAGAGGCCAC	948
XX					
PD	29-MAY-2003.	Qy	541	ACGCTTCTTCTCTCTTTGGGATCCTCTTTTCATCTGCTCTCTCTGCTGCTGCTCATGC	600
XX					
PF	18-JUL-2002; 2002US-00198846.	Db	949	ACGCTTCTTCTCTCTTTGGGATCCTCTTTTCATCTGCTCTCTCTGCTGCTGCTCATGC	1008
XX					
PR	18-JUL-2001; 2001US-0306220P.	Qy	601	TGTCAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTTTCCCTGTTGGTGAATTCGGG	660
XX					
PA	(MILL-) MILLENNIUM PHARM INC.	Db	1009	TGTCAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTTTCCCTGTTGGTGAATTCGGG	1068
XX					
PI	Lillie J, Xu Y, Wang Y, Steinmann K;	Qy	661	TCGGGCGTGGGCTTCAGCCCTAGTCAGGATGTTATCGCTATTGAAATATATATGTCGTAC	720
XX					
DR	WPI; 2003-787014/74.	Db	1069	TCGGGCGTGGGCTTCAGCCCTAGTCAGGATGTTATCGCTATTGAAATATATATGTCGTAC	1128
XX					
PT	Novel isolated polypeptide associated with breast cancer, useful for	Qy	721	CATGAATAGAACCAACGTCATCTCTTTCTGAGCTTTCGCTCCTCGTGCAGTGAAGA	780
XX					
PT	detecting presence of polypeptide in sample, as a marker for breast	Db	1129	CATGAATAGAACCAACGTCATCTCTTTCTGAGCTTTCGCTCCTCGTGCAGTGAAGA	1188
XX					
PT	cancer.	Qy	781	CTTTGTCTCTCTCTCACCTTACCTCTCTCTTTGATGGCGCTGACCTTCTCTCATGTCTCTC	840
XX					
PS	Disclosure; SEQ ID NO 10424; 36pp; English.	Db	1189	CTTTGTCTCTCTCTCACCTTACCTCTCTCTTTGATGGCGCTGACCTTCTCTCATGTCTCTC	1248
XX					
CC	The invention relates to an isolated polypeptide (I) associated with	Qy	841	CTTCACTCTCTGTGGTCTCTTACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC	900
XX					
CC	breast cancer which is encoded by a nucleic acid molecule comprising a	Db	1249	CTTCACTCTCTGTGGTCTCTTACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC	1308
XX					
CC	nucleotide sequence (SI). Further disclosed is an antibody that binds to	Qy	901	GATGCTCTCTCTCATTTGCCATCTGGGTGGGCTGGATCACCTCTCATGCTCTCTGACTT	960
XX					
CC	the polypeptide of the invention. The activity of the polypeptide of the	Db	1309	GATGCTCTCTCTCATTTGCCATCTGGGTGGGCTGGATCACCTCTCATGCTCTCTGACTT	1368
XX					
CC	invention may be described as cytostatic. The antibody is useful for	Qy	961	TGACCGCAGGTGGGATGACACCATCTCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGTGT	1020
XX					
CC	detecting the presence of (I) in a sample. Nucleic acid molecules of the	Db	1369	TGACCGCAGGTGGGATGACACCATCTCTCAGCTCCGCTTGGCTGGCCAAATGGCTGGTGT	1428
XX					
CC	invention are useful in the detection of breast tumours. (I) is useful as	Qy	1021	CTGTGTGGCTTATGTTAGTCCCGAGTTTGGCTGTCTCAAAAGCAACCAACCCCATGGA	1080
XX					
CC	a marker for breast cancer and in breast cancer therapy. Sequences given	Db	1429	CTGTGTGGCTTATGTTAGTCCCGAGTTTGGCTGTCTCAAAAGCAACCAACCCCATGGA	1488
XX					
CC	in records ACN78851-ACN92934 represent nucleic acid markers associated	Qy	1081	TTATCCGTGTGAGGATGCTTTCTGTAACCTCAACTCAACCTCAACCTCAACCTCAACCT	1140
XX					
CC	with breast cancer. Note: The sequence listing does not form part of the	Db	1489	TTATCCGTGTGAGGATGCTTTCTGTAACCTCAACTCAACCTCAACCTCAACCTCAACCT	1548
XX					
CC	specification but may be obtained in electronic format from the USPTO web	Qy	1141	GAACAGAGCTTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGACACGCT	1200
XX					
CC	site at seqdata.uspto.gov/sequence.html?DocID=2003099974	Db	1549	GAACAGAGCTTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGACACGCT	1608
XX					
SQ	Sequence 4239 BP; 859 A; 1161 C; 1075 G; 1141 T; 0 U; 3 Other;	Qy	1201	CTATGCCCTCTATTCACACATTTTTCAGCTGCAAGAACAGCCCTCCCAAGGAATTCCTC	1260
	Query Match 99.5%; Score 2443; DB 11; Length 4239;	Db	1609	CTATGCCCTCTATTCACACATTTTTCAGCTGCAAGAACAGCCCTCCCAAGGAATTCCTC	1668
	Best Local Similarity 100.0%; Pred. No. 0;	Qy	1261	CATCCCAAGGCGCCACGCTTGGCGGAGCCCTTCAAGAGCTATGAGTAAAGAAAGAGG	1320
	Matches 2443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Db	1669	CATCCCAAGGCGCCACGCTTGGCGGAGCCCTTCAAGAGCTATGAGTAAAGAAAGAGG	1728
Qy	1 ATACAGCATGAAGTCGCTGGAACTGGAATAGGCTGCTCTCTCCCTCGACCTCCCC	Qy	1321	CAGCTAACTCTGTCTGAAAGAGTGGGACAAATGACGCGCGGCGGAGATCTAGCGGAGC	1380
Db	409 ATAAAGCATGAAGTCGCTGGAACTGGAATAGGCTGCTCTCTCCCTCGACCTCCCC	Db	1729	CAGCTAACTCTGTCTGAAAGAGTGGGACAAATGACGCGCGGCGGAGATCTAGCGGAGC	1788
Qy	61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCTCCCTCGGCGAGGCGGCTTTATA	Qy	1381	TCAAGGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAAACCTGTACAGAACTGTACGGGA	1440
Db	469 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCTCCCTCGGCGAGGCGGCTTTATA	Db	1789	TCAAGGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAAACCTGTACAGAACTGTACGGGA	1848
Qy	121 ACAACTGCTCAGAGTGGAGGCGGGATAGCTGTCCAAGGTCTCCCGCAGCACTGAGGAG	Qy	1441	ACAGTTTGGCTCTCCCTCCAGCTCAACCAATTTCTTCCATGTGGGGCTGATGTGGGCT	1500
Db	529 ACAACTGCTCAGAGTGGAGGCGGGATAGCTGTCCAAGGTCTCCCGCAGCACTGAGGAG	Db	1849	ACAGTTTGGCTCTCCCTCCAGCTCAACCAATTTCTTCCATGTGGGGCTGATGTGGGCT	1908
Qy	181 CTGCGCTGCTGCTCTTGGCGGGGAGGAGCAAGTTCACGGCCAAAGCCTTGGC	Qy	1501	AGTAAGACTCCAGTCTTCTAGAGCGCTGTAGTATTTTTTTTTTTTTTTTTTTTGTCTCATCTTGG	1560
Db	589 CTCGCTGCTGCTCTTGGCGGGGAGGAGCAAGTTCACGGCCAAAGCCTTGGC	Db	1909	AGTAAGACTCCAGTCTTCTAGAGCGCTGTAGTATTTTTTTTTTTTTTTTTTTTGTCTCATCTTGG	1968
Qy	241 ACTAGGTCAGAAATGCTACAAAGTTCCTGATGTTGCGCAATGCGCTGAAATCCAA	Qy	1561	ATACTTTCTTTAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTCTTT	1620
Db	649 ACTAGGTCAGAAATGCTACAAAGTTCCTGATGTTGCGCAATGCGCTGAAATCCAA	Db	1969	ATACTTTCTTTAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTCTTT	2028
Qy	301 GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTTAGAAAACGTTGGCCAC				
Db	709 GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTTAGAAAACGTTGGCCAC				
Qy	361 AGCGGGGTTGTACCTCGGTGGCTTCATGCTCCTCCGATCTCTCGTCTGCAAGGT				
Db	769 AGCGGGGTTGTACCTCGGTGGCTTCATGCTCCTCCGATCTCTCGTCTGCAAGGT				
Qy	421 GCAGGACTCCAAAGCGGAAAATGCTGCTACTCAGTTTCTCTCTCTCTGCTGGTGTGT				
Db	829 GCAGGACTCCAAAGCGGAAAATGCTGCTACTCAGTTTCTCTCTCTCTCTGCTGGTGTGT				
Qy	481 GGGCATCTTTGGGCTCACCTTCGCTTTCATCATCGGACTGGAGGGAGCACAGGGCCAC				

QY 1621 GTTTTGTGAACAGGATCTTGCTCTGTCAACCCAGGCTTGAGTGCACTGGTGCAATCACAG 1680
DB GTTTTGTGAACAGGATCTTGCTCTGTCAACCCAGGCTTGAGTGCACTGGTGCAATCACAG 2088
QY 1681 CCCAGTGCAGCCTCGACCACTGTGCTCAGCAATCCTCCCATCTCCATCTCCCAAGTG 1740
DB CCCAGTGCAGCCTCGACCACTGTGCTCAGCAATCCTCCCATCTCCATCTCCCAAGTG 2148
QY 1741 CTGGGATGACAGCGCTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTTGCTGTTATTTT 1800
DB CTGGGATGACAGCGCTGAGCCACAGCTCCAGCCTAGGCCCTTAATCTTTGCTGTTATTTT 2208
QY 1801 CCATGAATAAAGGTCTGCTCATCTGAGCTCAGCTGAGCTCAGCAAGCTCTAGGGCCCTG 1860
DB CCATGAATAAAGGTCTGCTCATCTGAGCTCAGCTGAGCTCAGCAAGCTCTAGGGCCCTG 2268
QY 1861 CTCCTCTAACTCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
DB CTCCTCTAACTCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 2328
QY 1921 GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAAAC 1980
DB GAGCAAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAAAC 2388
QY 1981 TTGCTGGCACCCCGCTCCCCAAACCTTCTTGCTGGGTAGGAGGCTTAAAGATCACCC 2040
DB TTGCTGGCACCCCGCTCCCCAAACCTTCTTGCTGGGTAGGAGGCTTAAAGATCACCC 2448
QY 2041 TAAATTTACTCATCTCTAGTGTGCTGCCTCAGATTTGGGCTCAGAGCTCCCCAGCACCA 2100
DB TAAATTTACTCATCTCTAGTGTGCTGCCTCAGATTTGGGCTCAGAGCTCCCCAGCACCA 2508
QY 2101 ATTACAGGTACCCCTCTCTTCTGCACTGTCCCAAACTTGCTCAATTCGAGATC 2160
DB ATTACAGGTACCCCTCTCTTCTGCACTGTCCCAAACTTGCTCAATTCGAGATC 2568
QY 2161 TAATCTCCCTACGCTCTGCCAGGAATCTTTTCAGACCTCACTAGCAAGCCCGGTTG 2220
DB TAATCTCCCTACGCTCTGCCAGGAATCTTTTCAGACCTCACTAGCAAGCCCGGTTG 2628
QY 2221 CTCCTGTGAGGAGATTTGTAGATCATCTCACTCAATTCCTGGGCTGATCTCT 2280
DB CTCCTGTGAGGAGATTTGTAGATCATCTCACTCAATTCCTGGGCTGATCTCT 2688
QY 2281 CTCATCTTGACCCCAACCTCTGTAATAGATTTACCGCATTTACCGCTGCAATTCGTAA 2340
DB CTCATCTTGACCCCAACCTCTGTAATAGATTTACCGCATTTACCGCTGCAATTCGTAA 2748
QY 2341 GTGGGATGGTCTCTTAATGGAGGAGTGTTCATTTATTAATAGTTTATTCACCTGAGTAT 2400
DB GTGGGATGGTCTCTTAATGGAGGAGTGTTCATTTATTAATAGTTTATTCACCTGAGTAT 2808
QY 2401 GCAATTAAGATGTGGGCCACTCTTTTCATGTGGTGGCGAGCA 2443
DB GCAATTAAGATGTGGGCCACTCTTTTCATGTGGTGGCGAGCA 2851

RESULT 12
ID ADX97494
XX ADX97494 standard; DNA; 6730 BP.
XX AC
XX ADX97494;
DT 21-APR-2005 (first entry)
DE Pancreatic cancer associated human protein encoding DNA, SEQ ID 42.
XX pancreas tumor; cytostatic; gene; ds.
OS Homo sapiens.
XX.
PW EP1471075-A2.

XX PD 27-OCT-2004.
XX XX
PF 31-MAR-2004; 2004EP-00090124.
XX XX
PR 31-MAR-2003; 2003DE-01015834.
XX XX
PA (HINZ/) HINZMANN B.
PA (ROSE/) ROSENTHAL A.
PA (PILA/) PILARSKY C.
PA (DAHL/) DAHL E.
PA (SPEC/) SPECHT T.
PA (LICH/) LICHTNER R.
XX
PI Rosenthal A, Pilarsky C, Dahl E, Specht T, Bruemendorf T;
PI Lichtner R, Staub E, Roepcke S, Li X;
XX
XX WPI: 2004-768082/76.
DR P-PSDB; ADX97565.
XX
PT New nucleic acid differentially expressed in pancreatic tumor tissue, for
PT use as diagnostic agents and in screening for therapeutic agents.
XX
PS Claim 1; SEQ ID NO 42; 28pp; German.
XX
CC The invention relates to a novel human nucleic acid sequence of the
CC pancreas and its encoded protein. The invention further comprises:
CC proteins and peptides, preferably isolated, that contain a sequence
CC encoded by the novel nucleic acid; and methods for diagnosis and
CC treatment of pancreatic cancer, using a substance that inhibits or binds
CC to the protein or its DNA, including: an antisense oligonucleotide, short
CC interfering RNA or ribozyme directed against the pancreatic protein, an
CC organic molecule, particularly having a molecular weight below 5000,
CC especially 300, that binds to the pancreatic DNA, an aptamer or
CC (monoclonal) antibody, preferably human or humanized, that binds to the
CC pancreatic DNA, or an anti-idiotypic antibody raised against the
CC monoclonal antibody, any of which may be derivatized with a reporter
CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
CC human pancreatic proteins and their encoding DNA have cytostatic
CC activity. The novel sequences are useful for inhibiting transcription
CC and/or expression of genes and proteins associated with pancreatic
CC cancer. This polynucleotide sequence represents the DNA encoding one of
CC the novel human pancreatic proteins of the invention. Note: This sequence
CC is not shown in the specification, it has been electronically downloaded
CC from a DVD-rom provided with this specification by the European Patent
CC Office.
XX
SQ Sequence 6730 BP; 1469 A; 1763 C; 1555 G; 1943 T; 0 U; 0 Other;

Query Match 99.5%; Score 2443; DB 13; Length 6730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATACAGCATGAAGTCCGCTGGAATAGGCGTGCTCCCTCGACCCCTCCCC 60
DB 1 ATACAGCATGAAGTCCGCTGGAATAGGCGTGCTCCCTCGACCCCTCCCC 60
QY 61 TCCTTGCTCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGCGCGCTTTATA 120
DB 61 TCCTTGCTCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGCGCGCTTTATA 120
QY 121 ACAATGCTCAGAGTCCGAGGCGGATAGCTGTCCCAAGGTCTCCCGCAGACTGAGGAG 180
DB 121 ACAATGCTCAGAGTCCGAGGCGGATAGCTGTCCCAAGGTCTCCCGCAGACTGAGGAG 180
QY 181 CTCGCTGCTGCCCTCTTTCGCGCGGAGAGCAGACCAAGTTTCACGCGCAAGCCCTTGGC 240
DB 181 CTCGCTGCTGCCCTCTTTCGCGCGGAGAGCAGACCAAGTTTCACGCGCAAGCCCTTGGC 240
QY 241 ACTAGGGTCCAGAATGGCTTACAACTGCTCCCTGATGGTTGCCCAATGGCTGAAATCCAA 300
DB 241 ACTAGGGTCCAGAATGGCTTACAACTGCTCCCTGATGGTTGCCCAATGGCTGAAATCCAA 300

QY 301 GTACTCAGACTTGTGATGAAGCTGAAGCTGGGCGATCGTCTAGAAAACGGTGGCCAC 360
DB |||||
QY 301 GTACTCAGACTTGTGATGAAGCTGAAGCTGGGCGATCGTCTAGAAAACGGTGGCCAC 360
DB |||||
QY 361 AGCCGGGTTGTGACCTCGGTGGCTTTCATGCTCACTCTCCCGATCTCGTCTGCAAGGT 420
DB |||||
QY 361 AGCCGGGTTGTGACCTCGGTGGCTTTCATGCTCACTCTCCCGATCTCGTCTGCAAGGT 420
DB |||||
QY 421 GCAGGACTCCAAACGAGGAAAATGCTGCCTACTCAGTTCCTCTCTCTCGGTGTGTT 480
DB |||||
QY 421 GCAGGACTCCAAACGAGGAAAATGCTGCCTACTCAGTTCCTCTCTCTCGGTGTGTT 480
DB |||||
QY 481 GGGCATCTTTGGCCTCAGCTTCGCTTCATCATCGGACTGGAACGGGAGCACAGGGCCAC 540
DB |||||
QY 481 GGGCATCTTTGGCCTCAGCTTCGCTTCATCATCGGACTGGAACGGGAGCACAGGGCCAC 540
DB |||||
QY 541 ACGTTCCTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTCTGCTCATGC 600
DB |||||
QY 541 ACGTTCCTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTCTGCTCATGC 600
DB |||||
QY 601 TGTCACTCTGACCAAGCTCTCGGGGAGGAGCCCTTTCCCTGTTGGTGATCTCTGG 660
DB |||||
QY 601 TGTCACTCTGACCAAGCTCTCGGGGAGGAGCCCTTTCCCTGTTGGTGATCTCTGG 660
DB |||||
QY 661 TCTGGCCGTGGCTTCAAGCTTCAAGGATGTTATCGCTATTGAATATTGTCCTGAC 720
DB |||||
QY 661 TCTGGCCGTGGCTTCAAGCTTCAAGGATGTTATCGCTATTGAATATTGTCCTGAC 720
DB |||||
QY 721 CATGAATPAGGACCAACGTCAATGTCCTTTCTGAGCTTTCGCTCTCGTGCCTAATGAAGA 780
DB |||||
QY 721 CATGAATPAGGACCAACGTCAATGTCCTTTCTGAGCTTTCGCTCTCGTGCCTAATGAAGA 780
DB |||||
QY 781 CTTTGTCTCTGCTGCTACCTACCTCTCTCTCTGATGGCGCTGACCTTCTCATGCTCTC 840
DB |||||
QY 781 CTTTGTCTCTGCTGCTACCTACCTCTCTCTCTGATGGCGCTGACCTTCTCATGCTCTC 840
DB |||||
QY 841 CTTTCACTCTCTGCTGCTTCAAGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900
DB |||||
QY 841 CTTTCACTCTCTGCTGCTTCAAGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900
DB |||||
QY 901 GATGCTCTCTCAATGCTCAATGCTGGTGGCTGGATCACCTCTCATGCTTCTGACTT 960
DB |||||
QY 901 GATGCTCTCTCAATGCTCAATGCTGGTGGCTGGATCACCTCTCATGCTTCTGACTT 960
DB |||||
QY 961 TGACCGAGGTGGATGACACCACTCTCAGCTCGGCTTGGCTGCGCAATGGCTGGTGT 1020
DB |||||
QY 961 TGACCGAGGTGGATGACACCACTCTCAGCTCGGCTTGGCTGCGCAATGGCTGGTGT 1020
DB |||||
QY 1021 CTTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
DB |||||
QY 1021 CTTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
DB |||||
QY 1081 TTATCTGTTGAGGATGCTTCTGTAAACCTCAACTCGTGAAGAAGACTATGGTGTGGA 1140
DB |||||
QY 1081 TTATCTGTTGAGGATGCTTCTGTAAACCTCAACTCGTGAAGAAGACTATGGTGTGGA 1140
DB |||||
QY 1141 GAACAGAGCTCTCTCAAGAGAAATCACTCAAGTTTGAAGAGACAGGACAGCT 1200
DB |||||
QY 1141 GAACAGAGCTCTCTCAAGAGAAATCACTCAAGTTTGAAGAGACAGGACAGCT 1200
DB |||||
QY 1201 CTATGCCCCCTATTCCACACATTTTCACTGTCAGACACGCTCCCAAGAAATTTCTC 1260
DB |||||
QY 1201 CTATGCCCCCTATTCCACACATTTTCACTGTCAGACACGCTCCCAAGAAATTTCTC 1260
DB |||||
QY 1261 CATCCCAACGGGCCACGCTTGGCGAGCCCTTCAAAAGACTATGAAGTAAAGAAAGGG 1320
DB |||||
QY 1261 CATCCCAACGGGCCACGCTTGGCGAGCCCTTCAAAAGACTATGAAGTAAAGAAAGGG 1320
DB |||||
QY 1321 CAGCTAACTCTGCTGAGAGTGGGACAAATGAGCCGGCGGAGATCTAGCGGAGC 1380
DB |||||
QY 1321 CAGCTAACTCTGCTGAGAGTGGGACAAATGAGCCGGCGGAGATCTAGCGGAGC 1380
DB |||||
QY 1381 TCAAGGGATGTTGGGGAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA 1440

DB 1381 TCAAGGGATGTTGGGGAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA 1440
QY |||||
DB 1441 ACAGTTTGCCTCCCTCCAGCTCAACCAATTTCTCAATGTGGGGCTGATGTGGCT 1500
QY |||||
DB 1441 ACAGTTTGCCTCCCTCCAGCTCAACCAATTTCTCAATGTGGGGCTGATGTGGCT 1500
QY |||||
QY 1501 AGTAAGACTCCAGTCTTGAAGGGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTTTGG 1560
DB |||||
QY 1501 AGTAAGACTCCAGTCTTGAAGGGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTTTGG 1560
DB |||||
QY 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTGAAGCCCTTACTCTTTTCTTT 1620
DB |||||
QY 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTGAAGCCCTTACTCTTTTCTTT 1620
DB |||||
QY 1621 GTTTTGTGAAACAGGATCTTGTCTCACCAGGCTTGAAGTGTGAGTGTGCAATCAAG 1680
DB |||||
QY 1621 GTTTTGTGAAACAGGATCTTGTCTCACCAGGCTTGAAGTGTGAGTGTGCAATCAAG 1680
DB |||||
QY 1681 CCCAGTGCAGCTCCAGACCTGCTCAAGCAATCTCCCATCTCCCATCTCCCAAGTG 1740
DB |||||
QY 1681 CCCAGTGCAGCTCCAGACCTGCTCAAGCAATCTCCCATCTCCCATCTCCCAAGTG 1740
DB |||||
QY 1741 CTGGGATGACAGGCTGAGCCAGCTCCAGGCTTAGGGCTTAACTCTGCTGTTATTTT 1800
DB |||||
QY 1741 CTGGGATGACAGGCTGAGCCAGCTCCAGGCTTAGGGCTTAACTCTGCTGTTATTTT 1800
DB |||||
QY 1801 CCATGACTTAAAGGTCTGCTCAGCTCAGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1860
DB |||||
QY 1801 CCATGACTTAAAGGTCTGCTCAGCTCAGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1860
DB |||||
QY 1861 CTCTCTTAATCACTCAGCTGGGTTTGTGAGGCTCTGTGGCCAGGAGAGCTCATATCT 1920
DB |||||
QY 1861 CTCTCTTAATCACTCAGCTGGGTTTGTGAGGCTCTGTGGCCAGGAGAGCTCATATCT 1920
DB |||||
QY 1921 GAGCAAAAATAGCAAAAGGCTCTCTCAGGCCACTGGGCTTGAATCTTACACTTGAAGCCAAC 1980
DB |||||
QY 1921 GAGCAAAAATAGCAAAAGGCTCTCTCAGGCCACTGGGCTTGAATCTTACACTTGAAGCCAAC 1980
DB |||||
QY 1981 TTGCTGGACCCCGCTCCCAACCTTCTGCTGGTGGTGGAGAGGCTAAAGATCACCC 2040
DB |||||
QY 1981 TTGCTGGACCCCGCTCCCAACCTTCTGCTGGTGGTGGAGAGGCTAAAGATCACCC 2040
DB |||||
QY 2041 TAAATTTACTCATCTCTAGTCTGCTCACAATGGGCTCAGAGCTCCCGAGCACCA 2100
DB |||||
QY 2041 TAAATTTACTCATCTCTAGTCTGCTCACAATGGGCTCAGAGCTCCCGAGCACCA 2100
DB |||||
QY 2101 ATTCAGAGTCACTCTCTCTTGTGACTGTCCCAAACTTGTCTCAATTCGAGATC 2160
DB |||||
QY 2101 ATTCAGAGTCACTCTCTCTTGTGACTGTCCCAAACTTGTCTCAATTCGAGATC 2160
DB |||||
QY 2161 TAACTCCCGCTTACGCTCTGCGAGGAAATCTTTTCAAGCTCAGTACAGCAAGCCGGTGG 2220
DB |||||
QY 2161 TAACTCCCGCTTACGCTCTGCGAGGAAATCTTTTCAAGCTCAGTACAGCAAGCCGGTGG 2220
DB |||||
QY 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATCTGGGGCTGATCTCT 2280
DB |||||
QY 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATCTGGGGCTGATCTCT 2280
DB |||||
QY 2281 CTCTCTTGTGAGGAGAAATTTGTAGATCATTTTACCGCAATTTACCGCTGCAATTTCTGTA 2340
DB |||||
QY 2281 CTCTCTTGTGAGGAGAAATTTGTAGATCATTTTACCGCAATTTACCGCTGCAATTTCTGTA 2340
DB |||||
QY 2341 GTGGGATGCTCTCTTAAATGAGGAGTGTTCATGTATATAATAGTTATTACCTGAGTAT 2400
DB |||||
QY 2341 GTGGGATGCTCTCTTAAATGAGGAGTGTTCATGTATATAATAGTTATTACCTGAGTAT 2400
DB |||||
QY 2401 GCAATAAAGTGTGGTGGCCACTCTTTTCTGTTGGTGGGAGCA 2443
DB |||||
QY 2401 GCAATAAAGTGTGGTGGCCACTCTTTTCTGTTGGTGGGAGCA 2443
DB |||||

1141 GAACAGAGCCTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGAGACAGGGCACGCT 1200
1201 CTATGCCCCCTATTCACACATTTTCAGCTGCGAGACCGCTCCCAAAAGGAATTC 1260
1201 CTATGCCCCCTATTCACACATTTTCAGCTGCGAGACCGCTCCCAAAAGGAATTC 1260
1261 CATCCCAAGGGCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG 1320
1261 CATCCCAAGGGCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG 1320
1321 CAGCTAACTCTGCTCTGAGAGTGGGACAAATCAGCCGGCGGAGATCTAGCGGAGC 1380
1321 CAGCTAACTCTGCTCTGAGAGTGGGACAAATCAGCCGGCGGAGATCTAGCGGAGC 1380
1381 TCAAGGGATGTGGGCGAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
1381 TCAAGGGATGTGGGCGAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
1441 ACAGTTTGCCTCCCTCCAGCCTCAACCAAAATCTTCCATGCTGGGGCTGATGTGGCT 1500
1441 ACAGTTTGCCTCCCTCCAGCCTCAACCAAAATCTTCCATGCTGGGGCTGATGTGGCT 1500
1501 AGTAAGCTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
1501 AGTAAGCTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTTTTGT 1620
1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTTTTGT 1620
1621 GTTTTTTGAACAGGATCTTGTCTGTCAACCAAGGCTTGAAGTGGGCGGATCACAG 1680
1621 GTTTTTTGAACAGGATCTTGTCTGTCAACCAAGGCTTGAAGTGGGCGGATCACAG 1680
1681 CCAGTGCAGCCTCGACCACTGTCTGTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
1681 CCAGTGCAGCCTCGACCACTGTCTGTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
1741 CTGGGATGACAGGCGTGAGCCAGCTCCAGCCTAGGCGCTTAATCTTGTCTTTT 1800
1741 CTGGGATGACAGGCGTGAGCCAGCTCCAGCCTAGGCGCTTAATCTTGTCTTTT 1800
1801 CCATGCACTAAAGTCTTGTGCTATCTGAGCTCAGCTGAGTGCACAGCTCTAGGGCGCTG 1860
1801 GAGCAAAATTAAGTCTGTGCTATCTGAGCTCAGCTGAGTGCACAGCTCTAGGGCGCTG 1860
1861 CTCCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGCTCATATCT 1920
1861 CTCCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGCTCATATCT 1920
1921 GAGCAAAATTAAGCAAGGCTCTCAGCCCACTGGGCTGAATCTACCTGGAAGCAAC 1980
1921 GAGCAAAATTAAGCAAGGCTCTCAGCCCACTGGGCTGAATCTACCTGGAAGCAAC 1980
1981 TTGCTGGCACCCCGCTCCCAACCTCTTGTGCTGGTAGGAGGCTAAAGATCACCC 2040
1981 TTGCTGGCACCCCGCTCCCAACCTCTTGTGCTGGTAGGAGGCTAAAGATCACCC 2040
2041 TAAATTTACTCATCTCTCTAGTCTGCTCAGCCCACTGGGCTGAATCTACCTGGAAGCAAC 2100
2041 TAAATTTACTCATCTCTCTAGTCTGCTCAGCCCACTGGGCTGAATCTACCTGGAAGCAAC 2100
2101 ATTACAGGTCACCCCTCTCTTGTGCACTGTCCCAAACTTGTCTCAATTCGGAGATC 2160
2101 ATTACAGGTCACCCCTCTCTTGTGCACTGTCCCAAACTTGTCTCAATTCGGAGATC 2160
2161 TAATCTCCCTTACGCTCTGCGAGGAATCTTTTTCAGACCTCTAGCAAGCCCGGTTG 2220
2161 TAATCTCCCTTACGCTCTGCGAGGAATCTTTTTCAGACCTCTAGCAAGCCCGGTTG 2220
2221 CTCCTTGTGAGGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
2221 CTCCTTGTGAGGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280

2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTACGGATTACGGCTCATTTCTGTAA 2340
2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTACGGATTACGGCTCATTTCTGTAA 2340
2341 GTGGGCATGTCTCTTAATGGAGGAGTGTTCATTGTATATAAGTTATTACCTGAGTAT 2400
2341 GTGGGCATGTCTCTTAATGGAGGAGTGTTCATTGTATATAAGTTATTACCTGAGTAT 2400
2401 GCATTAAGATGTGTGGGCACCTCTTTCATGTGTGGTGGCAGCAAAAAA 2456
2401 GCAATAAGATGTGTGGGCACCTCTTTCATGTGTGGTGGCAGCAAAAAA 2456

RESULT 14

AAH14688
ID AAH14688 standard; cDNA; 2446 BP.

XX AAH14688;

DT 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:12388.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

PN 07-FEB-2001.

XX 28-JUL-2000; 2000BP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR Claim 8; SEQ ID NO 12388; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the 5602 complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAH13629 to AAH13632 represent CC oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 2446 BP; 507 A; 723 C; 572 G; 644 T; 0 U; 0 Other;

Query Match 99.3%; Score 2439.8; DB 4; Length 2446;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATAACAGCATGAAGTCGCCCTGGAACTGGAATAGGCGTGTCTCCCTCGACCCCTCCCC	60
DB	1	ATAACAGCATGAAGTCGCCCTGGAACTGGAATAGGCGTGTCTCCCTCGACCCCTCCCC	60
QY	61	TCCTTGTCCCTCTGCTCACCCTCGCTCGTTCCTCCCTCCGCGAGGCGCCGCTTTATA	120
DB	61	TCCTTGTCCCTCTGCTCACCCTCGCTCGTTCCTCCCTCCGCGAGGCGCCGCTTTATA	120
QY	121	ACAACTGCTCAGAGTCGAGGGCGGATAGTGTCAAGTGTCCCCAGCACTGAGGAG	180
DB	121	ACAACTGCTCAGAGTCGAGGGCGGATAGTGTCAAGTGTCCCCAGCACTGAGGAG	180
QY	181	CTCGCTGCTGCCCTCTTCGCGCGGGAAGCAGCACCAAGTTACGGCCAAAGCCTTGGC	240
DB	181	CTCGCTGCTGCCCTCTTCGCGCGGGAAGCAGCACCAAGTTACGGCCAAAGCCTTGGC	240
QY	241	ACTAGGGTCCAGAAATGGCTACAACAGTCCCTGATGTTGCGCAATGGCTGAAATCCAA	300
DB	241	ACTAGGGTCCAGAAATGGCTACAACAGTCCCTGATGTTGCGCAATGGCTGAAATCCAA	300
QY	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATCGCTAGAAACGGTGGCCAC	360
DB	301	GTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATCGCTAGAAACGGTGGCCAC	360
QY	361	AGCGGGTGTGACCTCGCTCGCTTCACTGCTCCGATCTCGTCCGATCTCGTCCGAGCT	420
DB	361	AGCGGGTGTGACCTCGCTCGCTTCACTGCTCCGATCTCGTCCGATCTCGTCCGAGCT	420
QY	421	GCAGACTCCAAACAGCGGAAATGCTGCTACTCAAGTTTCTCTCTCTCTCTCTCTCT	480
DB	421	GCAGACTCCAAACAGCGGAAATGCTGCTACTCAAGTTTCTCTCTCTCTCTCTCTCT	480
QY	481	GGGCACTTTGGGCTCACTTGGCTTCACTGCTCGGAGTGGAGGAGGAGGAGGAGGAG	540
DB	481	GGGCACTTTGGGCTCACTTGGCTTCACTGCTCGGAGTGGAGGAGGAGGAGGAGGAG	540
QY	541	AGCTTCTTCT	600
DB	541	AGCTTCTTCT	600
QY	601	TGTCAAGTCTGACCAAGCTCGTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	660
DB	601	TGTCAAGTCTGACCAAGCTCGTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	660
QY	661	TCTGGCGTGGGCTTCAAGCTAGTCAAGGATGTTATCGCTATTTGAATATATTTGCTG	720
DB	661	TCTGGCGTGGGCTTCAAGCTAGTCAAGGATGTTATCGCTATTTGAATATATTTGCTG	720
QY	721	CATGAATAGCAACAGCTCAATGCTCTTTCTGAGCTTTCGGCTCTCGTGGCAATGAAG	780
DB	721	CATGAATAGCAACAGCTCAATGCTCTTTCTGAGCTTTCGGCTCTCGTGGCAATGAAG	780
QY	781	CTTTGTCCTCTGCTCACTTCACTTCTTTGATGGGCTGACCTTCTCATGTCCTC	840
DB	781	CTTTGTCCTCTGCTCACTTCACTTCTTTGATGGGCTGACCTTCTCATGTCCTC	840
QY	841	CTTCACTTCTGCTGCTTCTTCACTGCTGGAAGACATGAGGAGGAGGAGGAGGAGGAG	900
DB	841	CTTCACTTCTGCTGCTTCTTCACTGCTGGAAGACATGAGGAGGAGGAGGAGGAGGAG	900
QY	901	GATGCTCTCTCATTGGGCTGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960
DB	901	GATGCTCTCTCATTGGGCTGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960

QY	961	TCACCGCAGTCGGATGACACCATCTCAGCTCCGCTTGGCTGCAATGCTGGTGTGTT	1020
DB	961	TCACCGCAGTCGGATGACACCATCTCAGCTCCGCTTGGCTGCAATGCTGGTGTGTT	1020
QY	1021	CCTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTCAAAAGCAACCAATGGA	1080
DB	1021	CCTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTCAAAAGCAACCAATGGA	1080
QY	1081	TTATCTCTGTTAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGCTGGA	1140
DB	1081	TTATCTCTGTTAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGCTGGA	1140
QY	1141	GAACGAGCTACTCTCAAGAGGAAATCACTCAAGTTTTGAAGAGAGGAGACAGCT	1200
DB	1141	GAACGAGCTACTCTCAAGAGGAAATCACTCAAGTTTTGAAGAGAGGAGACAGCT	1200
QY	1201	CTATGCCCTTATTCACACATTTTTCAGTGCAGAACCAAGCTTCCCAAGAGAAATCTC	1260
DB	1201	CTATGCCCTTATTCACACATTTTTCAGTGCAGAACCAAGCTTCCCAAGAGAAATCTC	1260
QY	1261	CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAGAGGG	1320
DB	1261	CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAGAGGG	1320
QY	1321	CAGCTAACTCTCTGCTGAAGAGTGGGCAAAATGTCAGCCGCGGCGAGACTAGCGGAGC	1380
DB	1321	CAGCTAACTCTCTGCTGAAGAGTGGGCAAAATGTCAGCCGCGGCGAGACTAGCGGAGC	1380
QY	1381	TCAAAGGATGTTGGCGGAATCTTGAGTCTTCTGAGAACTGTCAAGCACTACGGGA	1440
DB	1381	TCAAAGGATGTTGGCGGAATCTTGAGTCTTCTGAGAACTGTCAAGCACTACGGGA	1440
QY	1441	ACAGTTTGGCTCCCTCCACAGCTCAACCAATTTCTTCCATGCTGGGGCTGATGTTGGCT	1500
DB	1441	ACAGTTTGGCTCCCTCCACAGCTCAACCAATTTCTTCCATGCTGGGGCTGATGTTGGCT	1500
QY	1501	AGTAAGACTCCAGTTCTTAGAGGCTGATGTTATTTTTTTTTTTTTTTTTTTTTTTT	1560
DB	1501	AGTAAGACTCCAGTTCTTAGAGGCTGATGTTATTTTTTTTTTTTTTTTTTTTTTTT	1560
QY	1561	ATACCTTTTAAAGTGGGAGTCTCAGGCACTCAGTTTATAGCCCTTCTTTTGTGTTT	1620
DB	1561	ATACCTTTTAAAGTGGGAGTCTCAGGCACTCAGTTTATAGCCCTTCTTTTGTGTTT	1620
QY	1621	GTATTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGCATCACAG	1680
DB	1621	GTATTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGCATCACAG	1680
QY	1681	CCAGTGCAGCTCGACACCTGTGCTCAAGCAATCTCCCATCTCCCAAGTG	1740
DB	1681	CCAGTGCAGCTCGACACCTGTGCTCAAGCAATCTCCCATCTCCCAAGTG	1740
QY	1741	CTGGGATGACAGCGTGGAGCCACAGCTCCAGCTAGGCCCTTAATCTTGTCTTATTTT	1800
DB	1741	CTGGGATGACAGCGTGGAGCCACAGCTCCAGCTAGGCCCTTAATCTTGTCTTATTTT	1800
QY	1801	CCATGGCTTAAAGTCTGCTCATCTGAGCTCAGCTGGCTTCAACAGCTCTAGGGCCCTG	1860
DB	1801	CCATGGCTTAAAGTCTGCTCATCTGAGCTCAGCTGGCTTCAACAGCTCTAGGGCCCTG	1860
QY	1861	CTCCTCTAACTCAGTGGGTTTGTGAGCTCTGTGGCCAGAGCAGACCTCATATCT	1920
DB	1861	CTCCTCTAACTCAGTGGGTTTGTGAGCTCTGTGGCCAGAGCAGACCTCATATCT	1920
QY	1921	GAGCAAAATAGCAAAAGCTCTCTCAGCCCTGCGCTGAATCTACTACCTGGAAGCAAC	1980
DB	1921	GAGCAAAATAGCAAAAGCTCTCTCAGCCCTGCGCTGAATCTACTACCTGGAAGCAAC	1980
QY	1981	TTGCTGGCACCCCGCTCCCAACCTTCTTGTGGTGGGAGAGGCTAAAGATCACCC	2040
DB	1981	TTGCTGGCACCCCGCTCCCAACCTTCTTGTGGTGGGAGAGGCTAAAGATCACCC	2040

QY 2041 TAAATTTACTCATCTCTCTAGTGTGCTCACATTTGGGCTCAGAGCTCCCGACACCA 2100
DB |||||
QY 2041 TAAATTTACTCATCTCTCTAGTGTGCTCACATTTGGGCTCAGAGCTCCCGACACCA 2100
DB |||||
QY 2101 ATTACAGGTCACCCCTCTCTTTGCACTGTGCTCCCAAACTTGTGTCATTTCCGAGATC 2160
DB |||||
QY 2101 ATTACAGGTCACCCCTCTCTTTGCACTGTGCTCCCAAACTTGTGTCATTTCCGAGATC 2160
DB |||||
QY 2161 TAATCTCCCTCAGGCTGCTGCGAGGATTTCTTTCAGACCTCCTAGGACAGGCGCGTTG 2220
DB |||||
QY 2161 TAATCTCCCTCAGGCTGCTGCGAGGATTTCTTTCAGACCTCCTAGGACAGGCGCGTTG 2220
DB |||||
QY 2221 CTCCTTTGTCAGGAGATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
DB |||||
QY 2221 CTCCTTTGTCAGGAGATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
DB |||||
QY 2281 CTCATCTTGACACCCCACTCTCTTAATAGATTTACCGCATTTACCGCTGTCATTTCTGTAA 2340
DB |||||
QY 2281 CTCATCTTGACACCCCACTCTCTTAATAGATTTACCGCATTTACCGCTGTCATTTCTGTAA 2340
DB |||||
QY 2341 GTGGGATGCTCTCCTTAATGGAGGAGTGTTCATTTGATATTAATTAATTAATTAATTAAT 2400
DB |||||
QY 2341 GTGGGATGCTCTCCTTAATGGAGGAGTGTTCATTTGATATTAATTAATTAATTAATTAAT 2400
DB |||||
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATTTGTTGGTGGCAGCA 2443
DB |||||
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATTTGTTGGTGGCAGCA 2443
DB |||||

RESULT 15
ADR43992

ID ADR43992 standard; DNA; 2456 BP.
XX
AC ADR43992;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human breast tumour associated gene clone-02 SEQ ID NO:2.
XX
KW cancer; tumour; breast cancer; lung cancer; colon cancer; kidney cancer;
KW antibody; cytostatic; gene therapy; human; breast tumour tissue;
KW breast tissue; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2004074506-A2.
XX
PD 02-SEP-2004.
XX
PF 30-JAN-2004; 2004WO-US001030.
XX
PR 13-FEB-2003; 2003US-0447900P.
XX
PP (MERG-) MERGEN LTD.
XX
PI Hu Q, Peng A, Liu B, Love JR, Hao X, Ren M, Sheng Z;
XX
DR WPI; 2004-635589/61.
XX
PT Diagnosing and treating a cancer, e.g. breast, lung, colon, or kidney
PT cancer, comprises detecting and blocking the over expression of a gene of
PT a protein found in breast, lung, colon, or kidney tissue.
XX
PS Claim 2; SEQ ID NO 2; 220pp; English.

CC The present invention describes a method for diagnosing and treating a
CC cancer, e.g. breast, lung, colon, or kidney cancer. The method comprises
CC detecting and blocking the over expression of a gene of a protein found
CC in breast, lung, colon, or kidney tissue. Also described is an antibody
CC or binding portion of an antibody that specifically binds a protein found
CC in breast tissue, lung tissue, colon tissue, or kidney tissue. The
CC antibody has cytostatic activity, and can be used in gene therapy. The
CC methods, antibodies, polynucleotides and polypeptides from the present

CC invention are useful for detecting, diagnosing, preventing and treating
CC cancer, e.g. breast, lung, colon, or kidney cancer. The present sequence
CC represents a nucleotide sequence given in the present invention, which is
CC over expressed in human breast tumour tissue. N.B. All 385 sequences
CC referenced in this patent are detailed in the US provisional application
CC SN 60/447,900, filed 02/13/2003. In this application only one sequence,
CC representing the longest sequence of each of the 65 clones is listed in
CC Appendix A (and given as SEQ ID NO:1 to 65 in the Sequence Listing).
XX

SQ Sequence 2456 BP; 517 A; 710 C; 567 G; 643 T; 0 U; 19 Other;

Query Match 99.2%; Score 2437; DB 13; Length 2456;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2437; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATAAAGCATGAAGTCCGTCGCTCAACCTCGCTCGTTCCTCCCTCGGAGGCGCGCTTTATA 120
DB |||||
QY 61 TCCTTGTCTCTCTCTCAACCTCGCTCGTTCCTCCCTCGGAGGCGCGCTTTATA 120
DB |||||
QY 61 TCCTTGTCTCTCTCTCAACCTCGCTCGTTCCTCCCTCGGAGGCGCGCTTTATA 120
DB |||||
QY 121 ACAACTGCTCAGAGTCGAGGCGGATAGTCTGTCGAAGGTCTCCCGACACTGAGGAG 180
DB |||||
QY 121 ACAACTGCTCAGAGTCGAGGCGGATAGTCTGTCGAAGGTCTCCCGACACTGAGGAG 180
DB |||||
QY 181 CTCGCTGTGCTCTCTTGGCGCGGAAAGCAGCACAAGTTTCAACGCAACGCTTTGGC 240
DB |||||
QY 181 CTCGCTGTGCTCTCTTGGCGCGGAAAGCAGCACAAGTTTCAACGCAACGCTTTGGC 240
DB |||||
QY 241 ACTAGGGTCCAGATGGCTTACACAGTCCCTGATGTTGGCGCAATGGCTTGAATCAA 300
DB |||||
QY 241 ACTAGGGTCCAGATGGCTTACACAGTCCCTGATGTTGGCGCAATGGCTTGAATCAA 300
DB |||||
QY 301 GTACTACAGACTTTGTGATAAGGCTCAAGCTTGGGCGATCGTCTAGAAAACGGTGGCCAC 360
DB |||||
QY 301 GTACTACAGACTTTGTGATAAGGCTCAAGCTTGGGCGATCGTCTAGAAAACGGTGGCCAC 360
DB |||||
QY 361 AGCGGGGTGTGACCTCGGTGCGCTTCACTCTCCGATCTCTCGCTGCTGCAAGGT 420
DB |||||
QY 361 AGCGGGGTGTGACCTCGGTGCGCTTCACTCTCCGATCTCTCGCTGCTGCAAGGT 420
DB |||||
QY 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTGATTTCTTCTCTCTGCTGCTGTT 480
DB |||||
QY 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTGATTTCTTCTCTCTGCTGCTGTT 480
DB |||||
QY 481 GGGCATCTTTGGCTCACCCTTCCCTTCACTCGGACTGGAGGAGGACAGGCGCCAC 540
DB |||||
QY 481 GGGCATCTTTGGCTCACCCTTCCCTTCACTCGGACTGGAGGAGGACAGGCGCCAC 540
DB |||||
QY 541 AGCTTCTTCTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTGCTCATGC 600
DB |||||
QY 541 AGCTTCTTCTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTGCTCATGC 600
DB |||||
QY 601 TGTCACTCTGACCAAGCTCGTGGGGGAGGAGGCGCTTCCCTTGGTGGTCTGGG 660
DB |||||
QY 601 TGTCACTCTGACCAAGCTCGTGGGGGAGGAGGCGCTTCCCTTGGTGGTCTGGG 660
DB |||||
QY 661 TCTGGCGGTGGGCTTCAAGCTAGTCCAGGATGTTATCGCTATTGAATATATTTCTCTGAC 720
DB |||||
QY 661 TCTGGCGGTGGGCTTCAAGCTAGTCCAGGATGTTATCGCTATTGAATATATTTCTCTGAC 720
DB |||||
QY 721 CATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTCCGCTCTCTGCTGCAATGAAGA 780
DB |||||
QY 721 CATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTCCGCTCTCTGCTGCAATGAAGA 780
DB |||||
QY 781 CTTTGTCTCTCTCTCACTAGTCTCTTCTTGTAGGCGCTGACCTTCTCTATGCTCTC 840
DB |||||
QY 781 CTTTGTCTCTCTCTCACTAGTCTCTTCTTGTAGGCGCTGACCTTCTCTATGCTCTC 840
DB |||||
QY 841 CTTTCACTCTCTCTGTTCTCTTCAAGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900
DB |||||

Db 841 CTTACCTCTCTGTTTCCTTACGGGCTGGAGAGACATGGGGCCACATCTACCTCAC 900
Qy 901 GATGCTCCTCTCCATTTGCCATCTGGGTGGCCCTGGATCACCTGCTCATGCTTCTCTEACTT 960
Db 901 GATGCTCCTCTCCATTTGCCATCTGGGTGGCCCTGGATCACCTGCTCATGCTTCTCTGACTT 960
Qy 961 TGAACGCGAGGTGGGATGACACCATCTCTAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT 1020
Db 961 TGAACGCGAGGTGGGATGACACCATCTCTAGCTCCGCTTGGCTGCCAATGGCTGGGTGTT 1020
Qy 1021 CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGACCAAGAAACCCCATGGA 1080
Db 1021 CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGACCAAGAAACCCCATGGA 1080
Qy 1081 TTATCCTGTTAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
Db 1081 TTATCCTGTTAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA 1140
Qy 1141 GAACAGAGCTACTCTCAAGAGGAATCACTCAAGTTTTGAAGACAGGGGACACGCT 1200
Db 1141 GAACAGAGCTACTCTCAAGAGGAATCACTCAAGTTTTGAAGACAGGGGACACGCT 1200
Qy 1201 CTATGCCCTTATTCACACATTTTCAGCTGCAGAACAGCCTCCCCAAAGAAATTTCTC 1260
Db 1201 CTATGCCCTTATTCACACATTTTCAGCTGCAGAACAGCCTCCCCAAAGAAATTTCTC 1260
Qy 1261 CATCCACAGGGCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG 1320
Db 1261 CATCCACAGGGCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG 1320
Qy 1321 CAGCTAACTCTGCTGAAGAGTGGGACAAATGCAGCCGGCGGCGAGCTATAGCGGGAGC 1380
Db 1321 CAGCTAACTCTGCTGAAGAGTGGGACAAATGCAGCCGGCGGCGAGCTATAGCGGGAGC 1380
Qy 1381 TCAAGAGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
Db 1381 TCAAGAGGATGTGGGCGAAATCTTGAGTCTTCTGAGAAACTGTACAAGACACTACGGGA 1440
Qy 1441 ACAGTTTGCTCCTCCACAGCTCAACCAATCTTCCATGCTGGGGCTGATGGGCT 1500
Db 1441 ACAGTTTGCTCCTCCACAGCTCAACCAATCTTCCATGCTGGGGCTGATGGGCT 1500
Qy 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCCTTTGG 1560
Db 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCCTTTGG 1560
Qy 1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTGT 1620
Db 1561 ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTGT 1620
Qy 1621 GTTTTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGGTGGATCACAG 1680
Db 1621 GTTTTTTGAACAGGATCTTGTCTGTCAACCAGGCTTGAGTGCAGTGGTGGATCACAG 1680
Qy 1681 CCCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCCAAGTG 1740
Db 1681 CCCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCCAAGTG 1740
Qy 1741 CTGGGATGACAGGCTGAGCCACAGCTCCAGGCTAGGGCCCTTAATCTTGTGTTATTTT 1800
Db 1741 CTGGGATGACAGGCTGAGCCACAGCTCCAGGCTAGGGCCCTTAATCTTGTGTTATTTT 1800
Qy 1801 CCATGSACTAAAGGTCTGCTCATCTGAGCTCAAGTGGCTCAACAGCTCTAGGGCCCTG 1860
Db 1801 CCATGSACTAAAGGTCTGCTCATCTGAGCTCAAGTGGCTCAACAGCTCTAGGGCCCTG 1860
Qy 1861 CTCCTCTAACTCAGTGGGTTTTTGTAGGCTCTGTGGCCAGAGACCTGCTATCT 1920
Db 1861 CTCCTCTAACTCAGTGGGTTTTTGTAGGCTCTGTGGCCAGAGACCTGCTATCT 1920
Qy 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGGCCACTGGCCTGAATCTACACTGGAAGCCAAC 1980
Db 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGGCCACTGGCCTGAATCTACACTGGAAGCCAAC 1980

Qy 1981 TTGCTGGCACCCCGCTCCCCAAACCTTCTTGCTGGGTAGGAGAGGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCCGCTCCCCAAACCTTCTTGCTGGGTAGGAGAGGCTAAAGATCACCC 2040
Qy 2041 TAAATTTTACTCATCTCTCTAGTGTGCTCACAATGGGCTCAGCAGCTCCCCAGCACCA 2100
Db 2041 TAAATTTTACTCATCTCTCTAGTGTGCTCACAATGGGCTCAGCAGCTCCCCAGCACCA 2100
Qy 2101 ATTACAGGTCAACCCCTCTCTTTTGCACTGTCCCCAAACTTTGCTGTCAATTTCCGAGATC 2160
Db 2101 ATTACAGGTCAACCCCTCTCTTTTGCACTGTCCCCAAACTTTGCTGTCAATTTCCGAGATC 2160
Qy 2161 TAAATCTCCCTTACGCTCTGCCAGGAATTTCTTTTCAAGACCTCACTAGCACAAAGCCCGTTG 2220
Db 2161 TAAATCTCCCTTACGCTCTGCCAGGAATTTCTTTTCAAGACCTCACTAGCACAAAGCCCGTTG 2220
Qy 2221 CTCCTTGTGACAGGAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
Db 2221 CTCCTTGTGACAGGAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTTCT 2280
Qy 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTTTACCGCAATTTACGGCTGCATTTGTAA 2340
Db 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTTTACCGCAATTTACGGCTGCATTTGTAA 2340
Qy 2341 GTGGGATGGTCTCCTTAATGGAGGAGTGTTCATTTGTATTAATAGTTTATTTCACTGATAT 2400
Db 2341 GTGGGATGGTCTCCTTAATGGAGGAGTGTTCATTTGTATTAATAGTTTATTTCACTGATAT 2400
Qy 2401 GCAATAAAGATGTGTGGCCACTCTTTTCAATGTGTGGCAGCAAAAAA 2456
Db 2401 GCAATAAAGATGTGTGGCCACTCTTTTCAATGTGTGGCAGCAAAAAA 2456

Search completed: June 10, 2006, 15:19:12
Job time : 1410 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2006, 10:23:50 ; Search time 200 Seconds
(without alignments)
816.131 Million cell updates/sec

Title: US-10-600-816-3

Perfect score: 1865

Sequence: 1 MATTVPDGGCRNLGKSKYRL.....PRAHWPSPYKDYEYKKEGS 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq.8.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

9: geneseqp2005s.*

10: geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1865	100.0	357	4	AAB68891 Human REC
2	1865	100.0	357	4	AAB93311 Human pro
3	1865	100.0	357	6	ABR58569 Human can
4	1865	100.0	357	6	ABJ37054 Human bre
5	1865	100.0	357	6	ABR42649 Human GPC
6	1865	100.0	357	6	ABP81984 Human G
7	1865	100.0	357	7	ADD93239 RAIG1.1/
8	1865	100.0	357	7	ADN93302 Cancer/an
9	1865	100.0	357	8	ADI28460 Human GPC
10	1865	100.0	357	8	ADO29632 Human GPC
11	1865	100.0	357	8	ADR48222 Human ret
12	1865	100.0	357	8	ABM81354 Tumour-as
13	1865	100.0	357	8	ADU06613 Novel bro
14	1865	100.0	357	9	AEA00166 Human TAT
15	1865	100.0	357	9	AEA00190 Human TAT
16	1865	100.0	357	9	AEA00686 Human TAT
17	1865	100.0	357	9	AEA00710 Human TAT
18	1865	100.0	357	9	AED47479 Retinoic
19	1865	100.0	401	8	ADX97565 Pancreati
20	1865	100.0	409	5	ABP41877 Human ova
21	1865	100.0	595	7	ADF70471 Orphan re
22	1861	99.8	357	8	ADI28536 Human GPC
23	1861	99.8	357	8	ADI28535 Human GPC

ALIGNMENTS

RESULT 1

AAB68891

ID AAB68891 standard; protein; 357 AA.

XX

AC AAB68891;

DT 24-APR-2001 (first entry)

XX Human RECAP polypeptide, SEQ ID NO: 21.

XX

Human; RECAP; receptors and associated proteins; cerebroprotective; neurotropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV; antidiabetic; immunostimulant; immunomodulator; antiinflammatory; antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic; cytostatic; antibacterial; virucide; fungicide; protozoacide; antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer.

XX Homo sapiens.

XX WO200107612-A2.

XX 01-FEB-2001.

XX 21-JUL-2000; 2000WO-US020035.

XX 21-JUL-1999; 99US-0145232P.

XX 07-OCT-1999; 99US-0158578P.

XX 12-NOV-1999; 99US-0165192P.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;

XX Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;

XX WPI: 2001-168554/17.

XX N-PSDB; AAF58615.

XX Novel receptors and associated proteins for diagnosis and treatment of neurological disorders, immunological disorders including autoimmune/inflammatory disorders and cell proliferative disorders such as cancer.

XX Disclosure; Page 111-112; 128pp; English.

XX The present sequence is a human RECAP (receptors and associated proteins) polypeptide. RECAP polynucleotides and polypeptides are useful in the diagnosis, treatment and prevention of neurological disorders such as

CC stroke, Alzheimer's disease, Pick's disease, Huntington's disease,
 CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral
 CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD
 CC (Creutzfeldt-Jakob disease), GSS (Gerstmann-Strausler-Scheinker
 CC syndrome); immunological disorders, including autoimmune/inflammatory
 CC disorders such as AIDS, Digeorge's syndrome, severe combined
 CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
 CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
 CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,
 CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's disease, vital,
 CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
 CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis and cancer
 XX
 XX Sequence 357 AA;

Query Match 100.0%; Score 1865; DB 4; Length 357;
 Best Local Similarity 100.0%; Pred. No. 2.7e-203;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTVPDGCGRNLGKSKYYRLCDKAEAWGIVLTVATAGVVTSAFMTLPLILVCKVQDSN 60
 DB 1 MATTVPDGCGRNLGKSKYYRLCDKAEAWGIVLTVATAGVVTSAFMTLPLILVCKVQDSN 60
 QY 61 RRKMLPTQFLFLGLVGLGIFGLTFAFIIGLDGSTGPTFRFLFGILFSCFSCLLAHAVSLT 120
 DB 61 RRKMLPTQFLFLGLVGLGIFGLTFAFIIGLDGSTGPTFRFLFGILFSCFSCLLAHAVSLT 120
 QY 121 KLVGRKPLSLVILGLAVGSLVQDVIAEIVILTMNTNNVNFSELSAPRNEDFVLL 180
 DB 121 KLVGRKPLSLVILGLAVGSLVQDVIAEIVILTMNTNNVNFSELSAPRNEDFVLL 180
 QY 181 LTYVFLMALTFELMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLLMLPDPDRW 240
 DB 181 LTYVFLMALTFELMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLLMLPDPDRW 240
 QY 241 DDTILSSALAANGWVFLLAYVSPFEMLLTKQRPNDYPVEDAFCKPQLVKKSYGVENRAY 300
 DB 241 DDTILSSALAANGWVFLLAYVSPFEMLLTKQRPNDYPVEDAFCKPQLVKKSYGVENRAY 300
 QY 301 SOEETITQGFETGDTLYAPYSTHFOQNOPPOKEFSIPRAHAWPSPYKDYEVKKEGS 357
 DB 301 SOEETITQGFETGDTLYAPYSTHFOQNOPPOKEFSIPRAHAWPSPYKDYEVKKEGS 357

RESULT 2
 AAB93311
 ID AAB93311 standard; protein; 357 AA.
 AC AAB93311;
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12389.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 XX 29-JUL-1999; 99JP-00248036.
 XX 27-AUG-1999; 99JP-00300253.
 XX 11-JAN-2000; 2000JP-00118776.
 XX 02-MAY-2000; 2000JP-00183767.
 XX 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 XX cDNAs.
 PS Claim 8; SEQ ID NO 12389; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC polynucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 XX Sequence 357 AA;

Query Match 100.0%; Score 1865; DB 4; Length 357;
 Best Local Similarity 100.0%; Pred. No. 2.7e-203;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTVPDGCGRNLGKSKYYRLCDKAEAWGIVLTVATAGVVTSAFMTLPLILVCKVQDSN 60
 DB 1 MATTVPDGCGRNLGKSKYYRLCDKAEAWGIVLTVATAGVVTSAFMTLPLILVCKVQDSN 60
 QY 61 RRKMLPTQFLFLGLVGLGIFGLTFAFIIGLDGSTGPTFRFLFGILFSCFSCLLAHAVSLT 120
 DB 61 RRKMLPTQFLFLGLVGLGIFGLTFAFIIGLDGSTGPTFRFLFGILFSCFSCLLAHAVSLT 120
 QY 121 KLVGRKPLSLVILGLAVGSLVQDVIAEIVILTMNTNNVNFSELSAPRNEDFVLL 180
 DB 121 KLVGRKPLSLVILGLAVGSLVQDVIAEIVILTMNTNNVNFSELSAPRNEDFVLL 180
 QY 181 LTYVFLMALTFELMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLLMLPDPDRW 240
 DB 181 LTYVFLMALTFELMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLLMLPDPDRW 240
 QY 241 DDTILSSALAANGWVFLLAYVSPFEMLLTKQRPNDYPVEDAFCKPQLVKKSYGVENRAY 300
 DB 241 DDTILSSALAANGWVFLLAYVSPFEMLLTKQRPNDYPVEDAFCKPQLVKKSYGVENRAY 300
 QY 301 SOEETITQGFETGDTLYAPYSTHFOQNOPPOKEFSIPRAHAWPSPYKDYEVKKEGS 357
 DB 301 SOEETITQGFETGDTLYAPYSTHFOQNOPPOKEFSIPRAHAWPSPYKDYEVKKEGS 357

RESULT 3
 ABR58569
 ID ABR58569 standard; protein; 357 AA.
 XX ABR58569;
 XX ABR58569;
 XX 09-JUL-2003 (first entry)
 DT

QY 181 LTYVLFLMALTFMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLMLPDPDRRW 240
DB 181 LTYVLFLMALTFMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLMLPDPDRRW 240
QY 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPNDYPVEDAFCKPOLVKKSYGVENRAY 300
DB 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPNDYPVEDAFCKPOLVKKSYGVENRAY 300
QY 301 SQEETITQGFETGDTLYAPYSTHFQLOQPPOKESIPRAHAWPSPYKDYEVKKEGS 357
DB 301 SQEETITQGFETGDTLYAPYSTHFQLOQPPOKESIPRAHAWPSPYKDYEVKKEGS 357

RESULT 5

ABR42649
ID ABR42649 standard; protein; 357 AA.

XX ABR42649;

AC ABR42649;

DT 26-AUG-2003 (first entry)

XX Human GPCR-like retinoic acid-induced gene 1 protein.

XX Human; retinoic acid-induced gene 1; RAIG1; feeding; fasting; GPCR;
KW receptor; G-protein coupled receptor; anorectic; antidiabetic;
KW antidepressant; immunomodulator; transgenic; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 4..8 /note= "phosphorylated by casein kinase II"
FT Modified-site 8..14 /note= "N-myristoylated"
FT Modified-site 38..43 /note= "N-myristoylated"
FT Modified-site 59..61 /note= "phosphorylated by protein kinase C"
FT Modified-site 80..86 /note= "N-myristoylated"
FT Modified-site 98..93 /note= "N-myristoylated"
FT Modified-site 102..107 /note= "N-myristoylated"
FT Modified-site 124..127 /note= "Amidated"
FT Modified-site 136..142 /note= "N-myristoylated"
FT Modified-site 158..161 /note= "N-glycosylated"
FT Modified-site 201..206 /note= "N-myristoylated"
FT Modified-site 301..304 /note= "phosphorylated by casein kinase II"

XX WO2003016553-A2.

XX 27-FEB-2003.

XX 20-AUG-2002; 2002WO-US026510.

XX 20-AUG-2001; 2001US-0313940P.

XX (GETH) GENENTECH INC.
XX (CURA-) CURAGEN CORP.

XX Lewin DA, Stewart TA;

XX WPI; 2003-278580/27.

XX N-PSDB; ACC58386.

XX New G-protein coupled receptor-like retinoic acid induced gene 1 (GPCR-

PT like RAIG1) polypeptide and gene, useful for diagnosing or treating
PT metabolic disorders, e.g. obesity, anorexia, cachexia or diabetes.
XX Disclosure; Page 19-20; 150pp; English.

XX The present sequence is that of human G-protein coupled receptor-like
CC retinoic acid induced gene 1 (GPCR-like RAIG1) protein. This is the human
CC homologue of murine GPCR-like RAIG1 (see ABR42648). The murine GPCR-like
CC RAIG1 gene was shown to be differentially regulated during fasting-
CC feeding cycles, with moderate induction early in fasting, down-regulation
CC with extended fasting and 4-fold up-regulation with feeding in recovery
CC from fasting. The differentially expressed gene, its mRNA, and the
CC encoded protein, can each be manipulated to detect and treat metabolic
CC disorders associated with up- or down-regulation of GPCR-like RAIG1
CC activity, such as obesity, anorexia, cachexia or diabetes

XX SQ Sequence 357 AA;

Query Match 100.0%; Score 1865; DB 6; Length 357;

Best Local Similarity 100.0%; Pred. No. 2.7e-203;

Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTVDCRNLGSKYYRLCDKAEAWGIVLETATAGWTSVAFMLTLPILVCKVQDSN 60

DB 1 MATTVDCRNLGSKYYRLCDKAEAWGIVLETATAGWTSVAFMLTLPILVCKVQDSN 60

QY 61 RRKMLPTQFLFLGLVIGIFGLTFAPFIIGLDGSTGPTREFLFGILFSCILLAHAVSLT 120

DB 61 RRKMLPTQFLFLGLVIGIFGLTFAPFIIGLDGSTGPTREFLFGILFSCILLAHAVSLT 120

QY 121 KLVGRKPLSLVILGLAVGSLVQDVIAEYIVLTMTNTNVNVSSELSAPRNEDFVLL 180

DB 121 KLVGRKPLSLVILGLAVGSLVQDVIAEYIVLTMTNTNVNVSSELSAPRNEDFVLL 180

QY 181 LTYVLFLMALTFMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLMLPDPDRRW 240

DB 181 LTYVLFLMALTFMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAWITLMLPDPDRRW 240

QY 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPNDYPVEDAFCKPOLVKKSYGVENRAY 300

DB 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPNDYPVEDAFCKPOLVKKSYGVENRAY 300

QY 301 SQEETITQGFETGDTLYAPYSTHFQLOQPPOKESIPRAHAWPSPYKDYEVKKEGS 357

DB 301 SQEETITQGFETGDTLYAPYSTHFQLOQPPOKESIPRAHAWPSPYKDYEVKKEGS 357

RESULT 6

ABP81984

ID ABP81984 standard; protein; 357 AA.

XX AC ABP81984;

XX 04-MAR-2003 (first entry)

XX Human G protein-coupled receptor RAIG1 protein SEQ ID NO:454.

XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.

XX Homo sapiens.

XX WO200261087-A2.

08-AUG-2002.
19-DEC-2001; 2001WO-US050107.
19-DEC-2000; 2000US-0257144P.
(LIFE-) LIFESPAN BIOSCIENCES INC.
Burmer GC, Roush CL, Brown JP;
WPI; 2003-046718/04.
N-PSDB; AB242832.
New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.
Disclosure; Fig 1; 523pp; English.
The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (1) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention
Sequence 357 AA;
Query Match
Best Local Similarity 100.0%; Score 1865; DB 6; Length 357;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATTPDGCGRGLKSKYRLCDKAEAGVILETVATAGVVTVAFMLTILPVCKVQDSN 60
DQ 1 MATTPDGCGRGLKSKYRLCDKAEAGVILETVATAGVVTVAFMLTILPVCKVQDSN 60
QY 61 RKMLPTQFLVLGVIGLFAFIIGDSTGTPRFFLGLFSICFSCLLAHAVSLT 120
DQ 61 RKMLPTQFLVLGVIGLFAFIIGDSTGTPRFFLGLFSICFSCLLAHAVSLT 120
QY 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEYVLTMTNTNVNVSFELSAPRNEDFVLL 180
DQ 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEYVLTMTNTNVNVSFELSAPRNEDFVLL 180
QY 181 LTVVLFMALTFMSSFTCGSGFTGKRHGAHIYLTMLLSIAIWAVITLMLPDRRW 240
DQ 181 LTVVLFMALTFMSSFTCGSGFTGKRHGAHIYLTMLLSIAIWAVITLMLPDRRW 240
QY 241 DDTILSSALAANGWPELLAYVSPFELLTKQNPMDYPVEDAFCKPOLVKKSYGVENRAY 300
DQ 241 DDTILSSALAANGWPELLAYVSPFELLTKQNPMDYPVEDAFCKPOLVKKSYGVENRAY 300
QY 301 SQBEITQGFETGDTLYAPYSTHFLQNPQPKQFSPRAHAWPSPYKDYEVKKEGS 357

Db 301 SQBEITQGFETGDTLYAPYSTHFLQNPQPKQFSPRAHAWPSPYKDYEVKKEGS 357
RESULT 7
ADD93239
ID ADD93239 standard; protein; 357 AA.
XX
AC ADD93239;
XX
DT 29-JAN-2004 (first entry)
XX
DE RAIG1.
XX
KW retinoic acid-inducible gene 1; RAIG1; orphan;
KW G-protein coupled receptor; GPCR; chromosome 12; GPCR5B; carcinoma;
KW vaccine; breast cancer; pancreatic cancer; lung cancer; liver cancer;
KW ovarian cancer; colon cancer; osteosarcoma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 160..172
FT Peptide /note="Tandem mass spectrum peptide"
FT Peptide 334..348
FT Peptide /note="MALDI mass spectra peptide"
FT Peptide 340..348
FT Peptide /note="Tandem mass spectrum peptide"
XX
PN WO2003087832-A2.
XX
PD 23-OCT-2003.
XX
PF 10-APR-2003; 2003WO-GB001587.
XX
PR 11-APR-2002; 2002GB-00008331.
PR 17-SEP-2002; 2002GB-00021538.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Terrett JA;
XX
DR WPI; 2003-845382/78.
DR N-PSDB; ADD93240.
XX
PT Screening, diagnosing and/or treating carcinoma, including breast,
PT pancreatic, lung, liver, ovarian and colon cancer by detecting the change
PT in expression or activity of an RAIG1 polypeptide or encoding nucleic
PT acid molecule.
XX
PS Claim 1; Fig 1; 43pp; English.
XX
CC This sequence represents retinoic acid-inducible gene 1 (RAIG1)
CC polypeptide. RAIG1 is an orphan G-protein coupled receptor (GPCR) located
CC on chromosome 12. RAIG1 shows a restricted expression pattern compared to
CC a related receptor, GPCR5B, which is widely expressed in peripheral and
CC central tissues. The RAIG1 cDNA and polypeptide sequence may be used in
CC the method of the invention for screening for and/or diagnosis of carcinoma
CC carcinoma in a subject, and/or monitoring the effectiveness of carcinoma
CC therapy. The method comprises detecting and/or quantifying in a
CC biological sample obtained from the subject an RAIG1 polypeptide and a
CC nucleic acid molecule. The RAIG1 polypeptide and nucleic acid molecule
CC are useful in the manufacture of a medicament for the treatment of
CC carcinoma, where the composition is a vaccine. An agent which interacts
CC with or causes change in the expression or activity of an RAIG1
CC polypeptide or nucleic acid molecule, is also useful in the manufacture
CC of a medicament for the treatment of carcinoma that is breast cancer,
CC pancreatic cancer, lung cancer, liver cancer, ovarian cancer, colon
CC cancer and/or osteosarcoma. They can also be used in the diagnosis and
CC screening of such carcinomas.
XX
SQ Sequence 357 AA;

Query Match	100.0%;	Score 1865;	DB 7;	Length 357;
Best Local Similarity	100.0%;	Pred. No. 2.7e-203;		
Matches 357;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MATTVPDGCRCNGLSKKYRLCDKAEAWGIVLETVATAGVVTSAFVMTLPLILVCKVQDSN	60	
DB	1	MATTVPDGCRCNGLSKKYRLCDKAEAWGIVLETVATAGVVTSAFVMTLPLILVCKVQDSN	60	
QY	61	RRKMLPTQFLFLGLVGLFGLTFAPFIIGLDGSTGTRFRLFGILFSCFSCLLAHAVSLT	120	
DB	61	RRKMLPTQFLFLGLVGLFGLTFAPFIIGLDGSTGTRFRLFGILFSCFSCLLAHAVSLT	120	
QY	121	KLVRGKPLSLVILGLAVGFSLVQDVIAEIVILTMRTNNVNFSELSAPRNEDFVLL	180	
DB	121	KLVRGKPLSLVILGLAVGFSLVQDVIAEIVILTMRTNNVNFSELSAPRNEDFVLL	180	
QY	181	LTYYVLFMLALTFLMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAVITLLMLPDPDRRW	240	
DB	181	LTYYVLFMLALTFLMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAVITLLMLPDPDRRW	240	
QY	241	DDTILSSALAANGWVFLLAYVSPFELTQKRNPMNDYPVEDAFCKPQLVKKSYGVENRAY	300	
DB	241	DDTILSSALAANGWVFLLAYVSPFELTQKRNPMNDYPVEDAFCKPQLVKKSYGVENRAY	300	
QY	301	SOEITQGFEEGTDTLYAPYSTHFOLQNOPPOKFSIPRAHAWPSPYKDYEVKKEGS	357	
DB	301	SOEITQGFEEGTDTLYAPYSTHFOLQNOPPOKFSIPRAHAWPSPYKDYEVKKEGS	357	
RESULT 9				
ADN39302				
ID	ADN39302	standard; protein; 357 AA.		
XX	AC	ADN39302;		
XX	DT	17-JUN-2004 (first entry)		
XX	DE	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:620.		
XX	KW	Human; differential expression; cancer; angiogenic disorder;		
XX	KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;		
XX	KW	inflammatory disease; autoimmune disease;		
XX	KW	retinal neovascularisation syndrome; scarring; uterine fibroid;		
XX	KW	detection; diagnosis; prognosis; drug screening; drug targeting;		
XX	KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;		
XX	KW	vulnerary; gene therapy; vaccine.		
XX	OS	Homo sapiens.		
XX	PN	WO2003042661-A2.		
XX	PD	22-MAY-2003.		
XX	PF	13-NOV-2002; 2002WO-US036810.		
XX	XX	13-NOV-2001; 2001US-0350666P.		
PR	PR	21-NOV-2001; 2001US-0332464P.		
PR	PR	29-NOV-2001; 2001US-0334393P.		
PR	PR	03-DEC-2001; 2001US-0335394P.		
PR	PR	14-DEC-2001; 2001US-0340376P.		
PR	PR	08-JAN-2002; 2002US-0347211P.		
PR	PR	10-JAN-2002; 2002US-0347349P.		
PR	PR	13-FEB-2002; 2002US-0355250P.		
PR	PR	20-FEB-2002; 2002US-0356714P.		
PR	PR	29-MAR-2002; 2002US-0359077P.		
PR	PR	04-APR-2002; 2002US-036809P.		
PR	PR	12-APR-2002; 2002US-0372246P.		
PK	PK	05-JUN-2002; 2002US-0386614P.		
PR	PR	16-JUL-2002; 2002US-0396839P.		
PR	PR	22-JUL-2002; 2002US-0397755P.		
PR	PR	22-JUL-2002; 2002US-0397845P.		
PR	PR	09-SEP-2002; 2002US-0409450P.		

XX	PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX	PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;		
XX	PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;		
XX	DR	WPI; 2003-468649/44.		
XX	DR	N-PSDB; ADN39301.		
XX	PT	Determining the presence or absence of a pathological cell in a patient,		
XX	PT	useful for diagnosing, prognosing or treating cancer, comprises detecting		
XX	PT	a nucleic acid in a biological sample.		
XX	PS	Claim 12; SEQ ID NO 620; 1385pp; English.		
XX	CC	The invention relates to nucleic acids and proteins (ADN39683-ADN40064)		
XX	CC	whose expression is upregulated or downregulated in specific cancers or		
XX	CC	other diseases such as angiogenic or fibrotic disorders, and to methods		
XX	CC	of determining the presence or absence of a pathological cell in a		
XX	CC	patient by detecting a nucleic acid at least 80% identical to those of		
XX	CC	the invention or by detecting a polypeptide of the invention. The		
XX	CC	invention also relates to expression vectors and host cells comprising a		
XX	CC	nucleic acid of the invention; antibodies which specifically bind a		
XX	CC	polypeptide of the invention; use of such antibodies for drug targeting;		
XX	CC	and methods of screening for modulators of activity or expression of the		
XX	CC	polypeptides and nucleic acids. The nucleic acids, polypeptides,		
XX	CC	antibodies and methods are useful for diagnosing, prognosing and treating		
XX	CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,		
XX	CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal		
XX	CC	neovascularisation syndromes, scarring and uterine fibroids. They may		
XX	CC	also be useful in wound healing and in contraception. The present		
XX	CC	sequence represents a polypeptide of the invention.		
XX	SQ	Sequence 357 AA;		
	Query Match	100.0%;	Score 1865;	DB 7; Length 357;
	Best Local Similarity	100.0%;	Pred. No. 2.7e-203;	
	Matches 357;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MATTVPDGCRCNGLSKKYRLCDKAEAWGIVLETVATAGVVTSAFVMTLPLILVCKVQDSN	60	
DB	1	MATTVPDGCRCNGLSKKYRLCDKAEAWGIVLETVATAGVVTSAFVMTLPLILVCKVQDSN	60	
QY	61	RRKMLPTQFLFLGLVGLFGLTFAPFIIGLDGSTGTRFRLFGILFSCFSCLLAHAVSLT	120	
DB	61	RRKMLPTQFLFLGLVGLFGLTFAPFIIGLDGSTGTRFRLFGILFSCFSCLLAHAVSLT	120	
QY	121	KLVRGKPLSLVILGLAVGFSLVQDVIAEIVILTMRTNNVNFSELSAPRNEDFVLL	180	
DB	121	KLVRGKPLSLVILGLAVGFSLVQDVIAEIVILTMRTNNVNFSELSAPRNEDFVLL	180	
QY	181	LTYYVLFMLALTFLMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAVITLLMLPDPDRRW	240	
DB	181	LTYYVLFMLALTFLMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAVITLLMLPDPDRRW	240	
QY	241	DDTILSSALAANGWVFLLAYVSPFELTQKRNPMNDYPVEDAFCKPQLVKKSYGVENRAY	300	
DB	241	DDTILSSALAANGWVFLLAYVSPFELTQKRNPMNDYPVEDAFCKPQLVKKSYGVENRAY	300	
QY	301	SOEITQGFEEGTDTLYAPYSTHFOLQNOPPOKFSIPRAHAWPSPYKDYEVKKEGS	357	
DB	301	SOEITQGFEEGTDTLYAPYSTHFOLQNOPPOKFSIPRAHAWPSPYKDYEVKKEGS	357	
RESULT 9				
AD128460				
ID	AD128460	standard; protein; 357 AA.		
XX	AC	AD128460;		
XX	DT	22-APR-2004 (first entry)		
XX	DE	Human GPCR retinoic acid induced 3 (RAI-3) protein.		

XX Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytotstatic; cardiant; antiallergic;
KW broncholytic; gene therapy.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Region 1..357
FT /note= "Polypeptide specifically referred to in Claim 5"
FT Region 90..104
FT /note= "Polypeptide specifically referred to in Claim 5"
FT Region 155..178
FT /note= "Polypeptide specifically referred to in Claim 5"
FT Region 232..248
FT /note= "Polypeptide specifically referred to in Claim 5"
FT Region 269..284
FT /note= "Polypeptide specifically referred to in Claim 5"
FT Region 285..304
FT /note= "Polypeptide specifically referred to in Claim 5"
FT Region 340..353
FT /note= "Polypeptide specifically referred to in Claim 5"
XX WO2004001060-A2.
XX 31-DEC-2003.
XX 20-JUN-2003; 2003WO-US019255.
XX 20-JUN-2002; 2002US-0390850P.
XX 29-AUG-2002; 2002US-0407006P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Whitney GS, Opitack G, Garulacan L, Ramanathan CS, Mckinnon M;
PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX WPI: 2004-090973/09.
DR N-PSDB; AD128459.
XX New nucleic acid molecule encoding a human G-protein coupled receptor,
PT RAI3, useful for preventing, treating or ameliorating chronic obstructive
PT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
PT of COPD.
XX Claim 5; SEQ ID NO 3; 301pp; English.
XX The present sequence is the protein sequence of a human G-protein coupled
CC receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
CC Proteomics methods were used to isolate cigarette smoke-inducible
CC tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
CC identified as being tyrosine phosphorylated and/or as being
CC associated/complexed with tyrosine phosphorylated proteins only in cells
CC that had been exposed to cigarette smoke. Since RAI-3 is primarily
CC expressed in lung tissue, and since cigarette smoke is a major causative
CC factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
CC novel cellular target for identifying modulators, e.g. agonists or
CC antagonists, useful for the treatment and/or prevention of COPD and
CC related disorders such as emphysema and chronic bronchitis. RAI-3
CC modulators, e.g. agonists and antagonists, especially antisense
CC compounds, can be used to treat COPD and other disorders and diseases
CC associated with regulation of NF-kB and/or its associated or interacting
CC signalling molecules. Single nucleotide polymorphisms (SNPs) detected in
CC the RAI-3 gene are useful for determining COPD association in
CC individuals. 3 Of the SNPs result in alterations in the RAI-3 amino acid
CC sequence. Of these, Ser118Gly and Thr182A are at positions which are not
CC conserved between human and murine RAI-3 sequences, whereas Gln307Arg
CC involves a conserved amino acid residue. RAI-3 nucleic acid molecules and
CC polypeptides are useful for preventing, treating or ameliorating
CC disorders related to aberrant GPCR signaling or cell cycle regulation,
CC pulmonary disorders, inflammatory lung disorders, COPD, the underlying
CC symptoms of COPD, COPD-related disorders or conditions, autoimmune
CC disorders, disorders related to hyperimmune activity, inflammatory

CC conditions, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, renal diseases, ischaemia-reperfusion injury,
CC heart disorders, disorders related to aberrant signal transduction,
CC proliferation disorders, cancers, HIV infection, asthma, cystic fibrosis,
CC pulmonary fibrosis, ulcerative colitis, cerebral infarct, myocardial
CC infarct, diabetic nephropathy, allergic rhinitis, Crohn's disease,
CC atherosclerosis, rheumatoid arthritis, inflammatory/autoimmune disorders,
CC glioblastoma, pulmonary small cell undifferentiated carcinoma, carcinoma
CC of the breast, colon, lung, ovary, pancreas, prostate, non-Hodgkin's
CC lymphoma, disorders associated with aberrant cell adhesion, I-CAM
CC function and/or regulation, E-selectin function and/or regulation, or
CC aberrant NF-kB function and/or regulation (all claimed).
XX Sequence 357 AA;
SQ
Query Match 100.0%; Score 1865; DB 8; Length 357;
Best Local Similarity 100.0%; Pred. No. 2.7e-203;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATTVPDGCGRNGLSKKYVRLCDKAEAWGIVLETATAGVVTSAFMTLTPILVCKVQDSN 60
DB 1 MATTVPDGCGRNGLSKKYVRLCDKAEAWGIVLETATAGVVTSAFMTLTPILVCKVQDSN 60
QY 61 RRKMLPTQFLFLGVLGIFGLTFPAFLIGLDGSGTGRFELFGILFSCSLAHAVSLT 120
DB 61 RRKMLPTQFLFLGVLGIFGLTFPAFLIGLDGSGTGRFELFGILFSCSLAHAVSLT 120
QY 121 KLVGRKPLSLVILGLAVGFSLVQDVIAIEYIVLTMTNNTNVNVSFELSAPRRNEDFVLL 180
DB 121 KLVGRKPLSLVILGLAVGFSLVQDVIAIEYIVLTMTNNTNVNVSFELSAPRRNEDFVLL 180
QY 181 LTYVFLMALTFMSSFTFCGSGFTGKRGHAIHYLTMLLSIAIWAIVITLLMLPDDRWR 240
DB 181 LTYVFLMALTFMSSFTFCGSGFTGKRGHAIHYLTMLLSIAIWAIVITLLMLPDDRWR 240
QY 241 DDTILSSALAANGWVFLLAYVSPFELWLLTKQRPMDYPVEDAFCKPQLVKKSIVGNRAY 300
DB 241 DDTILSSALAANGWVFLLAYVSPFELWLLTKQRPMDYPVEDAFCKPQLVKKSIVGNRAY 300
QY 301 SQEETQGFEEGTDLTAPYSTHFQIQNQPPQKEFSIPRAHAWPSYKDYEVKKEGS 357
DB 301 SQEETQGFEEGTDLTAPYSTHFQIQNQPPQKEFSIPRAHAWPSYKDYEVKKEGS 357
RESULT 10
ADO29632
ID ADO29632 standard; protein; 357 AA.
XX
AC ADO29632;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human GPCR RAI3, SEQ ID NO:734.
XX
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antineoplastic;
KW cytostatic; antiinflammatory; vasotropic; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
XX receptor.
XX Homo sapiens.
XX

PN WO2004040000-A2.
 XX 13-MAY-2004.
 XX
 XX
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 XX 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 XX (PRIM-) PRIMAL INC.
 PA
 XX Gaifanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
 XX
 DR WPI; 2004-390329/36.
 DR N-PSDB; ADO30035.
 XX
 XX Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 XX Claim 151; SEQ ID NO 734; 542pp; English.
 XX
 XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 357 AA;

Query Match 100.0%; Score 1865; DB 8; Length 357;
 Best Local Similarity 100.0%; Pred. No. 2.7e-203;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTPDGCGRNGLKSKYYRLCDKAEAWGIVLTETATAGVVTVAFMLTLPILVCKVQDSN 60
 DB |||||
 DB 1 MATTPDGCGRNGLKSKYYRLCDKAEAWGIVLTETATAGVVTVAFMLTLPILVCKVQDSN 60
 QY 61 RRKMLPTQFLVLGVGLFGLTFAFIIIGDGTGTPRFLFGLFSICFSCILAHAVSLT 120
 DB |||||
 DB 61 RRKMLPTQFLVLGVGLFGLTFAFIIIGDGTGTPRFLFGLFSICFSCILAHAVSLT 120
 QY 121 KLVRGKPLSLVILGLAVGFLSLVQDVIAIEIVILTMNTNNVNFSELSAPRNEDFVLL 180
 DB |||||
 DB 121 KLVRGKPLSLVILGLAVGFLSLVQDVIAIEIVILTMNTNNVNFSELSAPRNEDFVLL 180

QY 181 LTYVLFMLALTFLMSSFTFCGSFTGKXRGHAIYLTMLLSIAIWAWITLLMLPDFDRW 240
 DB |||||
 DB 181 LTYVLFMLALTFLMSSFTFCGSFTGKXRGHAIYLTMLLSIAIWAWITLLMLPDFDRW 240
 QY 241 DDTILSSALAANGWVFLLAYVSPFLLTKQNPMDYVVEDAFCKPOLVKKSYGVENRAY 300
 DB |||||
 DB 241 DDTILSSALAANGWVFLLAYVSPFLLTKQNPMDYVVEDAFCKPOLVKKSYGVENRAY 300
 QY 301 SOEITQGFEEGTDLTLYAPYSTHFOLQNOPPOKFSIPRAHAWPSPYKDYEVKKEGS 357
 DB |||||
 DB 301 SOEITQGFEEGTDLTLYAPYSTHFOLQNOPPOKFSIPRAHAWPSPYKDYEVKKEGS 357
 RESULT 11
 ADR48222
 ID ADR48222 standard; protein; 357 AA.
 XX
 AC ADR48222;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human retinoic acid induced 3 (RAI3) protein SEQ:10.
 XX
 KW pancreatic cancer-associated transcript; pancreatic cancer; human;
 KW cytostatic; gene therapy; protein therapy; retinoic acid induced 3; RAI3.
 XX
 OS Homo sapiens.
 XX
 PN WO2004074510-A1.
 XX
 PD 02-SEP-2004.
 XX
 PF 18-FEB-2004; 2004WO-AU000194.
 XX
 PR 18-FEB-2003; 2003AU-00900747.
 XX
 PA (GARV-) GARVAN INST MEDICAL RES.
 PI
 XX Biankin A, Segara D, Henshall S, Sutherland R;
 DR WPI; 2004-635591/61.
 DR N-PSDB; ADR48221.
 XX
 PT Detecting pancreatic cancer-associated transcript in a biological sample,
 PT useful for diagnosing or treating the disease, comprises contacting the
 PT sample with a polynucleotide that selectively hybridizes to a specific
 PT sequence.
 XX
 PS Claim 7; SEQ ID NO 10; 263pp; English.
 CC
 CC The present invention describes a method for detecting a pancreatic
 CC cancer-associated transcript in a biological sample. The method comprises
 CC contacting the biological sample with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to a sequence as shown in
 CC any one of Tables 3 to 25 in the specification or having the GenBank
 CC Accession Number AF279145. Also described: (1) diagnosing pancreatic
 CC cancer in a human or animal subject being tested, determining the
 CC likelihood that a subject having a pancreatic cancer will survive, or
 CC determining the suitability of a subject having a pancreatic cancer for
 CC surgical resection therapy; (2) detecting a pancreatic cancer-associated
 CC polypeptide in a biological sample; (3) determining the likelihood that a
 CC subject having a pancreatic cancer will survive; and (4) monitoring the
 CC efficacy of a therapeutic treatment of pancreatic cancer. A pancreatic
 CC cancer-associated transcript has cytostatic activity, and can be used in
 CC gene and protein therapy. A pancreatic cancer-associated transcript
 CC polynucleotide, a vector comprising the polynucleotide, an isolated
 CC polypeptide or an antibody that binds to the isolated polypeptide can be
 CC used for diagnosing or prognosing pancreatic cancer or for preparing a
 CC medicament for the treatment of pancreatic cancer. The prognostic or
 CC diagnostic methods are useful for the early detection of pancreatic
 CC cancer or its metastases, and for monitoring the progress of disease such
 CC as during remission or following surgery or chemotherapy. The present


```
XX 09-APR-2003; 2003DE-01016701.
XX 09-APR-2003; 2003DE-01016701.
XX (HINZ/) HINZMANN B.
PA (HERM/) HERMANN K.
PA (CAST/) HEIDEN CASTANOS-VELEZ E.
XX Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pillarsky C;
XX WPI; 2004-786403/78.
DR N-PSDB; ADU06126.
XX New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT cancer and in screening for therapeutic and diagnostic agents.
XX Claim 2; SEQ ID NO 839; 1381pp; German.
XX This invention relates to a novel isolated nucleic acid associated with
CC bronchial cancer comprising 489 defined sequences given in the
CC specification. The invention may be useful for the production of
CC compounds with a cytostatic activity through the inhibition of expression
CC or activity of tumour-associated proteins. The novel DNA sequences and
CC the proteins/peptides encoded by them are used for detecting bronchial
CC cancer or determining the risk of developing it and to screen for
CC specific binding partners of the DNA or protein sequences, where the
CC binding partners are potentially useful as agents for treating or
CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a protein
CC encoded by a novel bronchial cancer-associated human gene sequence of the
CC invention.
XX SQ Sequence 357 AA;
Query Match 100.0%; Score 1865; DB 8; Length 357;
Best Local Similarity 100.0%; Pred. No. 2.7e-203;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATTVPDGCGRNLKSKYYRLCDKAEAWGIVLETVATAGVVTSAFMLTLPILVCKVQDSN 60
Db |||||
1 MATTVPDGCGRNLKSKYYRLCDKAEAWGIVLETVATAGVVTSAFMLTLPILVCKVQDSN 60
QY 61 RKMLPTQFLFLGLVGLIFGLTFAFTIGDGTGTPTREFLGILFSCILAHAVSLT 120
Db |||||
61 RKMLPTQFLFLGLVGLIFGLTFAFTIGDGTGTPTREFLGILFSCILAHAVSLT 120
QY 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMNTNNVFSLSAPRNEDFVLL 180
Db |||||
121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMNTNNVFSLSAPRNEDFVLL 180
QY 181 LTYVLFMALTFMSSFTFCGSGFTGKRGAHIYLTMLLSIAIWAIVITLMLPFDRRW 240
Db |||||
181 LTYVLFMALTFMSSFTFCGSGFTGKRGAHIYLTMLLSIAIWAIVITLMLPFDRRW 240
QY 241 DDTILSSALAANGWFLIAYVSPFWLLTKORNPMDYPVEDAFCKPOLVKKSYGVENRAY 300
Db |||||
241 DDTILSSALAANGWFLIAYVSPFWLLTKORNPMDYPVEDAFCKPOLVKKSYGVENRAY 300
QY 301 SOEETITQFEETGDTLYAPYSTHFQIQNOPPOKFEFSIPRAHAWSPYKDYEVKKEGS 357
Db |||||
301 SOEETITQFEETGDTLYAPYSTHFQIQNOPPOKFEFSIPRAHAWSPYKDYEVKKEGS 357
RESULT 14
AEA00166
ID# AEA00166 standard; protein; 357 AA.
XX AEA00166;
```

```
XX 28-JUL-2005 (first entry)
XX Human TAT115 protein SEQ ID NO:118.
XX tumor-associated antigen; cytostatic; breast tumor; endocrine disease;
KW gynecology and obstetrics; neoplasm; colon tumor;
KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
KW ovary tumor; skin tumor; liver tumor.
XX Homo sapiens.
XX US2005106644-A1.
XX 19-MAY-2005.
XX 08-SEP-2004; 2004US-00936626.
XX 20-JUN-2001; 2001US-0299500P.
PR 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323268P.
PR 19-JUN-2002; 2002US-00177488.
PR 26-MAR-2004; 2004US-0557118P.
PR 04-AUG-2004; 2004US-0598899P.
XX (GETH ) GENENTECH INC.
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koeppen H, Phillips HS;
PI Polakowski P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sliwowski M;
XX WPI; 2005-384304/39.
DR N-PSDB; AEA00088.
XX Novel isolated antibody capable of binding to tumor-associated antigenic
PT target polypeptide, useful for treating cell proliferative disorder e.g.
PT cancer.
XX Claim 1; SEQ ID NO 118; 337pp; English.
XX The invention relates to a novel isolated antibody binding to a
CC polypeptide having at least 80% sequence identity to a polypeptide having
CC any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
CC target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
CC specification, a polypeptide having any one of (AEA00127-AEA00202),
CC lacking its associated signal peptide, or an extracellular domain of a
CC polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
CC encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
CC An antibody of the invention has cytostatic activity. The antibody is
CC useful for inhibiting growth of a cell expressing TAT188, which involves
CC contacting the cell with the antibody. The cell is a cancer cell chosen
CC from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
CC liver cell. The cancer cell is a mammalian cell, preferably a human cell.
CC The antibody is also useful for detecting the level of TAT188
CC polypeptide expressed in a test cell relative to a control cell, and for
CC detecting the level of TAT188 polypeptide or a polypeptide having at
CC least 80 % sequence identity to the TAT188 polypeptide sequence in a test
CC cell relative to a control cell. The antibody is useful for treating a
CC preventing a cell proliferative disorder associated with increased
CC expression or activity of a polypeptide having at least 80 % identity to
CC a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
CC The method of the invention is useful for inhibiting the growth of a
CC cancer cell. The present sequence represents a polypeptide of the
CC invention.
XX SQ Sequence 357 AA;
Query Match 100.0%; Score 1865; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 2.7e-203;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATTVPDGCGRNLKSKYYRLCDKAEAWGIVLETVATAGVVTSAFMLTLPILVCKVQDSN 60
```

Db 1 MATTVPGCRNGLSKYYRLCDKAEAGVILETVATAGVVTSAFVMTLPILVCKVQDSN 60
QY 61 RRKMLPTQFLFLLGVIGFGLTFAPFIIGLDGSTGTPRFFLFGILFSCSCLLAHAVSLT 120
Db 61 RRKMLPTQFLFLLGVIGFGLTFAPFIIGLDGSTGTPRFFLFGILFSCSCLLAHAVSLT 120
QY 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMNRTNNVSELSAPRRNEDFVLL 180
Db 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMNRTNNVSELSAPRRNEDFVLL 180
QY 181 LTYVFLMALTFILMSSFTFCGSFTGKRGHAIYLTMLLSIAIWWAITLLMLPDPDRW 240
Db 181 LTYVFLMALTFILMSSFTFCGSFTGKRGHAIYLTMLLSIAIWWAITLLMLPDPDRW 240
QY 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPMDYPVEDAFCKPQLVKKSYGVENRAY 300
Db 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPMDYPVEDAFCKPQLVKKSYGVENRAY 300
QY 301 SOEITQGFEEETGDTLYAPYSTHFOLQNPPOKESIPRAHAWPSYKDYEVKKEGS 357
Db 301 SOEITQGFEEETGDTLYAPYSTHFOLQNPPOKESIPRAHAWPSYKDYEVKKEGS 357

RESULT 15
AEA00190
ID AEA00190 standard; protein; 357 AA.
XX
AC AEA00190;
DT 28-JUL-2005 (first entry)
DE Human TAT143 protein SEQ ID NO:142.
XX
KW tumor-associated antigen; cytostatic; breast tumor; endocrine disease;
KW gynecology and obstetrics; neoplasm; colon tumor;
KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
KW ovary tumor; skin tumor; liver tumor.
XX
OS Homo sapiens.
XX
PN US2005106644-A1.
XX
PD 19-MAY-2005.
XX
PF 08-SEP-2004; 2004US-00936626.
XX
PR 20-JUN-2001; 2001US-0299500P.
PR 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323268P.
PR 19-JUN-2002; 2002US-00177488.
PR 26-MAR-2004; 2004US-0557116P.
PR 04-AUG-2004; 2004US-0598899P.
XX
PA (GETH) GENENTECH INC.
XX
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koeppen H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sliwkowski M;
XX
DR WPI; 2005-384304/39.
DR N-PSDB; AEA00112.
XX
XX Novel isolated antibody capable of binding to tumor-associated antigenic
PT target polypeptide, useful for treating cell proliferative disorder e.g.
PT cancer.
XX
PS Claim 1; SEQ ID NO 142; 337pp; English.
XX
XX The invention relates to a novel isolated antibody binding to a
CC polypeptide having at least 80% sequence identity to a polypeptide having
CC any one of 76 fully defined 182-910 amino acid tumor-associated antigenic

CC target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
CC specification, a polypeptide having any one of (AEA00127-AEA00202),
CC lacking its associated signal peptide, or an extracellular domain of a
CC polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
CC encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
CC An antibody of the invention has cytostatic activity. The antibody is
CC useful for inhibiting growth of a cell expressing TAT188, which involves
CC contacting the cell with the antibody. The cell is a cancer cell chosen
CC from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
CC liver cell. The cancer cell is a mammalian cell, preferably a human cell.
CC The antibody is also useful for detecting the level of TAT188
CC polypeptide expressed in a test cell relative to a control cell, and for
CC detecting the level of TAT188 polypeptide or a polypeptide having at
CC least 80 % sequence identity to the TAT188 polypeptide sequence in a test
CC cell relative to a control cell. The antibody is useful for treating a
CC preventing a cell proliferative disorder associated with increased
CC expression or activity of a polypeptide having at least 80 % identity to
CC a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
CC The method of the invention is useful for inhibiting the growth of a
CC cancer cell. The present sequence represents a polypeptide of the
CC invention.
XX
SQ Sequence 357 AA;
Query Match 100.0%; Score 1865; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 2.7e-203;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATTVPGCRNGLSKYYRLCDKAEAGVILETVATAGVVTSAFVMTLPILVCKVQDSN 60
Db 1 MATTVPGCRNGLSKYYRLCDKAEAGVILETVATAGVVTSAFVMTLPILVCKVQDSN 60
QY 61 RRKMLPTQFLFLLGVIGFGLTFAPFIIGLDGSTGTPRFFLFGILFSCSCLLAHAVSLT 120
Db 61 RRKMLPTQFLFLLGVIGFGLTFAPFIIGLDGSTGTPRFFLFGILFSCSCLLAHAVSLT 120
QY 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMNRTNNVSELSAPRRNEDFVLL 180
Db 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMNRTNNVSELSAPRRNEDFVLL 180
QY 181 LTYVFLMALTFILMSSFTFCGSFTGKRGHAIYLTMLLSIAIWWAITLLMLPDPDRW 240
Db 181 LTYVFLMALTFILMSSFTFCGSFTGKRGHAIYLTMLLSIAIWWAITLLMLPDPDRW 240
QY 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPMDYPVEDAFCKPQLVKKSYGVENRAY 300
Db 241 DDTILSSALAANGWVFLAYVSPFWLLTKQRPMDYPVEDAFCKPQLVKKSYGVENRAY 300
QY 301 SOEITQGFEEETGDTLYAPYSTHFOLQNPPOKESIPRAHAWPSYKDYEVKKEGS 357
Db 301 SOEITQGFEEETGDTLYAPYSTHFOLQNPPOKESIPRAHAWPSYKDYEVKKEGS 357
Search completed: June 8, 2006, 10:27:24
Job time : 204 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2006, 10:32:39 ; Search time 50 Seconds
(without alignments)
624.969 Million cell updates/sec

Title: US-10-600-816-3

Perfect score: 1865

Sequence: 1 MATTVPDGCNGLSKYRL.....PRAHWPSPYKDYEVKKEGS 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5 COMB.pdp:*
2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6 COMB.pdp:*
3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7 COMB.pdp:*
4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H COMB.pdp:*
5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS COMB.pdp:*
6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE COMB.pdp:*
7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1865	100.0	357	US-09-949-016-6443	Sequence 6443, Ap
2	1865	100.0	390	US-09-949-016-9484	Sequence 9484, Ap
3	1782	95.5	347	US-09-188-930-326	Sequence 326, App
4	1782	95.5	347	US-09-312-283C-326	Sequence 326, App
5	350	18.8	68	US-09-188-930-123	Sequence 123, App
6	350	18.8	68	US-09-312-283C-123	Sequence 123, App
7	157.5	8.4	256	US-09-964-956-85	Sequence 85, Appl
8	146.5	7.9	879	US-09-964-956-53	Sequence 53, Appl
9	146.5	7.9	879	US-09-964-956-54	Sequence 54, Appl
10	145.5	7.8	879	US-09-964-956-17	Sequence 17, Appl
11	140.5	7.5	879	US-09-964-956-52	Sequence 52, Appl
12	139.5	7.5	879	US-08-072-574-6	Sequence 6, Appl1
13	139.5	7.5	879	US-08-486-270-6	Sequence 6, Appl1
14	139.5	7.5	879	US-08-367-264-6	Sequence 6, Appl1
15	139.5	7.5	879	US-08-794-158-2	Sequence 2, Appl1
16	139.5	7.5	879	US-09-153-757-6	Sequence 6, Appl1
17	139.5	7.5	879	US-09-459-715-6	Sequence 6, Appl1
18	136	7.3	200	US-09-305-258-349	Sequence 349, App
19	136	7.3	200	US-10-004-860-349	Sequence 349, App
20	131.5	7.1	872	US-08-337-797A-2	Sequence 2, Appl1
21	131.5	7.1	872	US-09-258-523-2	Sequence 2, Appl1
22	131	7.0	877	US-09-619-353-2	Sequence 2, Appl1
23	124	6.6	863	US-09-619-353-14	Sequence 14, Appl
24	119.5	6.4	1078	US-10-125-772-28	Sequence 28, Appl
25	119.5	6.4	1078	US-10-125-772-28	Sequence 28, Appl
26	119.5	6.4	1078	US-10-125-772-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-6443
; Sequence 6443, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6443

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Human

; US-09-949-016-6443

Query Match 100.0%; Score 1865; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 3.2e-194;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATTVPDGCNGLSKYRLCDKAEAWGIVLETVATAGVVTVA	FMFLTPILVCKVQDSN	60
Db	1	MATTVPDGCNGLSKYRLCDKAEAWGIVLETVATAGVVTVA	FMFLTPILVCKVQDSN	60
QY	61	RRKMLPTQFLFLLGVIGFGLTFATFIIGDSTGTPRFLFG	ILFSCFLLAHAVSLT	120
Db	61	RRKMLPTQFLFLLGVIGFGLTFATFIIGDSTGTPRFLFG	ILFSCFLLAHAVSLT	120
QY	121	KLVRGRKPLSLAVILGLAVGFSLVODVIAIEYIVLTNR	TNNVNFSELAPRNEFDVLL	180
Db	121	KLVRGRKPLSLAVILGLAVGFSLVODVIAIEYIVLTNR	TNNVNFSELAPRNEFDVLL	180
QY	181	LTIVLFMLALTFLMSSFTFCGSFTGKRGHAIYLTMLLS	IAIWAWITLLMLPDFDRRW	240
Db	181	LTIVLFMLALTFLMSSFTFCGSFTGKRGHAIYLTMLLS	IAIWAWITLLMLPDFDRRW	240
QY	241	DDTILSSALAANGWVFLAYVSPFWLLTKQRNPMDPYVE	DAFCKPOLVKKSIVENRAY	300
Db	241	DDTILSSALAANGWVFLAYVSPFWLLTKQRNPMDPYVE	DAFCKPOLVKKSIVENRAY	300

```
QY 301 SOEETQGFETGDTLYAPYSTHFOQNPPQKFSIPRAHAWPSPYKDYEVKKEGS 357
|||||
Db 301 SOEETQGFETGDTLYAPYSTHFOQNPPQKFSIPRAHAWPSPYKDYEVKKEGS 357

RESULT 2
US-09-949-016-9484
; Sequence 9484, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2000-04-14
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9484
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9484

Query Match 100.0%; Score 1865; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.7e-194;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTPDGCNGLKSKYRLCDKAEAGVILETVATAGVVTSAVAFMLTLPILVCKVQDSN 60
|||||
Db 34 MATTPDGCNGLKSKYRLCDKAEAGVILETVATAGVVTSAVAFMLTLPILVCKVQDSN 93

QY 61 RRKMLPTQFLFLGLVGLTFAFIIGLDGSTGPTFRFLFGILFISICFSCILAHAVSLT 120
|||||
Db 94 RRKMLPTQFLFLGLVGLTFAFIIGLDGSTGPTFRFLFGILFISICFSCILAHAVSLT 153

QY 121 KLVGRKPLSLVILGLAVGFLVQDVIAIEYIVLTMTNTNVNVSSELSAPRNEDFVLL 180
|||||
Db 154 KLVGRKPLSLVILGLAVGFLVQDVIAIEYIVLTMTNTNVNVSSELSAPRNEDFVLL 213

QY 181 LTYVLFMLALTFMSSFTFCGSGFTGKRGHAIYLTMLLSIAIWAIVITLMLPDPDRW 240
|||||
Db 214 LTYVLFMLALTFMSSFTFCGSGFTGKRGHAIYLTMLLSIAIWAIVITLMLPDPDRW 273

QY 241 DDTILSSALAANGWFLLAYVSPFLLTKORNPMDYPVEDAFCKPQLVKKSYGVENRAY 300
|||||
Db 274 DDTILSSALAANGWFLLAYVSPFLLTKORNPMDYPVEDAFCKPQLVKKSYGVENRAY 333

QY 301 SOEETQGFETGDTLYAPYSTHFOQNPPQKFSIPRAHAWPSPYKDYEVKKEGS 357
|||||
Db 334 SOEETQGFETGDTLYAPYSTHFOQNPPQKFSIPRAHAWPSPYKDYEVKKEGS 390

RESULT 3
US-09-188-930-326
; Sequence 326, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
```

```
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 326
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Human
US-09-188-930-326

Query Match 95.5%; Score 1782; DB 2; Length 347;
Best Local Similarity 99.1%; Pred. No. 3.3e-185;
Matches 341; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 KSKYYRLCDKAEAGVILETVATAGVVTSAVAFMLTLPILVCKVQDSNRRKMLPTQFLFLL 73
: : |||||
Db 4 RPPYRLCDKAEAGVILETVATAGVVTSAVAFMLTLPILVCKVQDSNRRKMLPTQFLFLL 63

QY 74 GVLGIFGLTFAFIIGLDGSTGPTFRFLFGILFISICFSCILAHAVSLTKLVGRKPLSLV 133
|||||
Db 64 GVLGIFGLTFAFIIGLDGSTGPTFRFLFGILFISICFSCILAHAVSLTKLVGRKPLSLV 123

QY 134 ILGLAVGFLVQDVIAIEYIVLTMTNTNVNVSSELSAPRNEDFVLLTYVLFMALTFEL 193
|||||
Db 124 ILGLAVGFLVQDVIAIEYIVLTMTNTNVNVSSELSAPRNEDFVLLTYVLFMALTFEL 183

QY 194 MSSFTFCGSGFTGKRGHAIYLTMLLSIAIWAIVITLMLPDPDRWDITILSSALAANG 253
|||||
Db 184 MSSFTFCGSGFTGKRGHAIYLTMLLSIAIWAIVITLMLPDPDRWDITILSSALAANG 243

QY 254 WFLLAYVSPFLLTKORNPMDYPVEDAFCKPQLVKKSYGVENRAYSOEETQGFETG 313
|||||
Db 244 WFLLAYVSPFLLTKORNPMDYPVEDAFCKPQLVKKSYGVENRAYSOEETQGFETG 303

QY 314 DTLYAPYSTHFOQNPPQKFSIPRAHAWPSPYKDYEVKKEGS 357
|||||
Db 304 DTLYAPYSTHFOQNPPQKFSIPRAHAWPSPYKDYEVKKEGS 347

RESULT 4
US-09-312-283C-326
; Sequence 326, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 326
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-326

Query Match 95.5%; Score 1782; DB 2; Length 347;
Best Local Similarity 99.1%; Pred. No. 3.3e-185;
Matches 341; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 KSKYYRLCDKAEAGVILETVATAGVVTSAVAFMLTLPILVCKVQDSNRRKMLPTQFLFLL 73
: : |||||
Db 4 RPPYRLCDKAEAGVILETVATAGVVTSAVAFMLTLPILVCKVQDSNRRKMLPTQFLFLL 63

QY 74 GVLGIFGLTFAFIIGLDGSTGPTFRFLFGILFISICFSCILAHAVSLTKLVGRKPLSLV 133
|||||
Db 64 GVLGIFGLTFAFIIGLDGSTGPTFRFLFGILFISICFSCILAHAVSLTKLVGRKPLSLV 123
```


QY 134 ILGLAVGFSIVODVIAIEYIVLTMMNTNNVSELSAPRNEDFVLLLTVVLFMALTEL 193
DB 124 ILGLAVGFSIVODVIAIEYIVLTMMNTNNVSELSAPRNEDFVLLLTVVLFMALTEL 183
QY 194 MSSFTFCGSGFTGKRGHAIYLTMLLSIAIIVAWITLLMLPDRDRWDDTILSSALAANG 253
DB 184 MSSFTFCGSGFTGKRGHAIYLTMLLSIAIIVAWITLLMLPDRDRWDDTILSSALAANG 243
QY 254 WFLLAYVSPFELLTKQRNPMDYPVEDAFCKPQLVKYSYGVENRAYSOEETIQGPEETG 313
DB 244 WFLLAYVSPFELLTKQRNPMDYPVEDAFCKPQLVKYSYGVENRAYSOEETIQGPEETG 303
QY 314 DTLYAPYSTHFQLONPQPKESIPRAHAWPSPYKDYEVKKEGS 357
DB 304 DTLYAPYSTHFQLONPQPKESIPRAHAWPSPYKDYEVKKEGS 347

RESULT 5

US-09-188-930-123

; Sequence 123, Application US/09188930A

; Patent No. 6150502

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James Greg

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; TITLE OF INVENTION: and Methods For Their Use

; FILE REFERENCE: 11000.1011c1

; CURRENT APPLICATION NUMBER: US/09/188,930A

; CURRENT FILING DATE: 1998-11-09

; NUMBER OF SEQ ID NOS: 348

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 123

; LENGTH: 68

; TYPE: PRT

; ORGANISM: Human

US-09-188-930-123

Query Match 18.8%; Score 350; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 2e-30;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 MLTLPILVCKVQDSNRRKMLPTQFLFLGLVIGLTFAPFIIGLDGSTGPTTRFFLFGILF 105
DB 1 MLTLPILVCKVQDSNRRKMLPTQFLFLGLVIGLTFAPFIIGLDGSTGPTTRFFLFGILF 60

QY 106 SICFSCLL 113

DB 61 SICFSCLL 68

RESULT 6

US-09-312-283C-123

; Sequence 123, Application US/09312283C

; Patent No. 6573095

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 123

; LENGTH: 68
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-123

Query Match 18.8%; Score 350; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 2e-30;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 MLTLPILVCKVQDSNRRKMLPTQFLFLGLVIGLTFAPFIIGLDGSTGPTTRFFLFGILF 105
DB 1 MLTLPILVCKVQDSNRRKMLPTQFLFLGLVIGLTFAPFIIGLDGSTGPTTRFFLFGILF 60

QY 106 SICFSCLL 113

DB 61 SICFSCLL 68

RESULT 7

US-09-964-956-85

; Sequence 85, Application US/09964956

; Patent No. 6875570

; GENERAL INFORMATION:

; APPLICANT: Gerlach, Valerie L

; APPLICANT: MacDougall, John R

; APPLICANT: Smithson, Glennda

; APPLICANT: Millet, Isabelle

; APPLICANT: Stone, David

; APPLICANT: Gunther, Erik

; APPLICANT: Ellerman, Karen

; APPLICANT: Grosse, William M

; APPLICANT: Alsobrook II, John P

; APPLICANT: Lepley, Denise M

; APPLICANT: Burgess, Catherine E

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Leach, Martin D

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. 6875570el Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-124

; CURRENT APPLICATION NUMBER: US/09/964,956

; CURRENT FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/235,631

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/235,633

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/235,808

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,064

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,065

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,066

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/236,135

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: 60/237,434

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/238,321

; PRIOR FILING DATE: 2000-10-05

; PRIOR APPLICATION NUMBER: 60/238,399

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,396

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/276,667

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/294,823

; PRIOR FILING DATE: 2001-05-31

; PRIOR APPLICATION NUMBER: 60/304,868

; PRIOR FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 127

; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 85
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 7
; OTHER INFORMATION: transmembrane receptor (metabotropic E family)
; OTHER INFORMATION: Consensus Sequence
US-09-964-956-85

Query Match
Best Local Similarity 8.4%; Score 157.5; DB 2; Length 256;
Matches 65; Conservative 43; Mismatches 75; Indels 53; Gaps 11;

QY 28 GIVLETATAGVVTVAFMLTLPILVCKVQD-----SNRRKMLPTQFLFLLGLVGLIF 79
DQ 2 GIVLVALAVLGI-----LTLFVLVVFVKHRDTPIVKASRE---LSYLLGLIGILCY 51
QY 80 GLTFAFIIGDSTGPT-----RFFLEGLIFSCFCLLAHAVSLTKLVGRKPLS----- 130
DQ 52 LCSFLFI-----GKPSSTCILRIILFGLGFTLCYSALLAKTRVLRIFRAKPKGSKPKF 107
QY 131 -----LLVILGLAVGFSVQDVIAEYIVLTMRNTNNVFSE-----LSAPRRN-EDPV 178
DQ 108 ISFWAQLVILV-----LIQVILCVMLVVEPPPTDIYSEKEKILCECKGSGWAPV 163
QY 179 LLTLYVFLMALTFMLSSFT--FCGSFTGKRGHGAHIYLTMLLSIAIWMVAMITLIM 232
DQ 164 VVLGVDGLLAVLCTFLAPLTRNLNPFNEAK---PIGFSMLTFCIVWVAFPIVL 215

RESULT 8
US-09-964-956-53
; Sequence 53, Application US/09964956
; Patent No. 6875570
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. 6875570el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964,956
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321

; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276,667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294,823
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304,868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 53
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-964-956-53

Query Match
Best Local Similarity 7.9%; Score 146.5; DB 2; Length 879;
Matches 73; Conservative 57; Mismatches 148; Indels 59; Gaps 11;

QY 8 GCRNGLKSKYVRLCDKAEAWGIVLETATAGVVTVAFMLTLPILVCKVQDSNR----- 61
DQ 561 GCYN-LPBDYIRWED---AWAIGPVITACLG-----FMCTCIVITVFIKHNTPLVKAS 610
QY 62 -RKMLPTQFLFLLGLVGLIFGLTFAFIIGDSTGPTFRFLFGILFSCFCLLAHAVSLT 120
DQ 611 GRELC---YILLFGVLSYCWTFFFIAKPSVICALRRLGLGTSFAICYSALLTKNCIA 667
QY 121 KLVGRK-----PLSLVILGLAVGFSVQDVIAEYIVLTMRNTNNVFSE----- 167
DQ 668 RIFDGVKNGAQRKPKFISPSQVFI---CIGLILVQIVMVVWLILETPTGTRRYTLPEKRE 724
QY 168 ---LSAPRRNEDFVLLTY--VLFMLALTFLMSSFTFCGSFTGKRGHGAHIYLTMLLSIA 222
DQ 725 TVILKCNVKSMLISLTYDVVVLCTVYAFKTKCPENFNEAKPIGTMTTTCI----- 780
QY 223 IWVAMITLMLPDPFDRRWDITLSSALAANGVVFLLAVVSPFELLTKORNPMDYPVEDA 282
DQ 781 IWLAFPLFIYVTSDDYRVQTTTMCISVSLSGFPVLGCLFAPKVHIVLFQ----- 829
QY 283 FCKPOLVKSGYVENRAYSQBEITQGFETGDTLYAP 319
DQ 830 ---PQKNVVTTRLHLNRFVSGTATTYSQSSASTYVP 863

RESULT 9
US-09-964-956-54
; Sequence 54, Application US/09964956
; Patent No. 6875570
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. 6875570el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964,956
; CURRENT FILING DATE: 2001-09-26
```

;; PRIOR APPLICATION NUMBER: 60/235,631
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/235,633
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/235,808
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,064
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,065
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,066
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,135
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: 60/237,434
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/238,321
;; PRIOR FILING DATE: 2000-10-05
;; PRIOR APPLICATION NUMBER: 60/238,399
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/238,396
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/276,667
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/294,823
;; PRIOR FILING DATE: 2001-05-31
;; PRIOR APPLICATION NUMBER: 60/304,868
;; PRIOR FILING DATE: 2001-07-12
;; NUMBER OF SEQ ID NOS: 127
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 54
;; LENGTH: 879
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-964-956-54

Query Match 7.9%; Score 146.5; DB 2; Length 879;
Best Local Similarity 21.7%; Pred. No. 1.1e-06;
Matches 73; Conservative 57; Mismatches 148; Indels 59; Gaps 11;

QY 8 GCRNGLSKYRCLDKAEANGIVLETATAGVVTSAFMTLTPILVCKVQDSNR----- 61
DB 561 GCYN-LPEDYIRWED---AWAIGPVTIACLG-----FMCTCIVITVFIKHNTPLVKAS 610

QY 62 -RKMLPTQFLFLGLVGIPLTFAFIIGLDGSTGPTFRFLFGILFSCISCLLAHAVSLT 120
DB 611 GRELC---YILLFGVSLSYCMTEFFFAKPSVICALRLRLGLGTSFAICYSALLTKTNCA 667

QY 121 KLVGRGK-----PLSLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNFSE--- 167
DB 668 RIFDGVNGAQRPKFISPSQVFI---CLGLILVQIVMVSVMLILETPGTRRYTLPEKRE 724

QY 168 ---LSAPRNEDFVLLTY--VLFMLALTFLMSSFTFCGSGFTGKRGHGAHIYLTMLLSIA 222
DB 725 TVILKCNVNDSSMLISLTYDVVLVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCI--- 780

QY 223 IIVAWITLMLPDRDRDDTLISSALANGWVFLAYSPFELWLTQKRNPMYDVEDA 282
DB 781 IWLAFILFIYVTSDDYRVQTTMCISVLSGFFVLGCLFAPKVIHLFQ----- 829

QY 283 FCKPOLVKSGYVENRAYSQEITQGFETGDTLYAP 319
DB 830 ---PQKNVVTHRLHNRFSVSGTATTYSQSSASTYVP 863

RESULT 10
US-09-964-956-17
; Sequence 17, Application US/09964956
; Patent No. 6875570
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda

;; APPLICANT: Millet, Isabelle
;; APPLICANT: Stone, David
;; APPLICANT: Gunther, Erik
;; APPLICANT: Ellerman, Karen
;; APPLICANT: Grosse, William M
;; APPLICANT: Alsbrook II, John P
;; APPLICANT: Lepley, Denise M
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Kekuda, Ramesh
;; APPLICANT: Spytek, Kimberly A
;; APPLICANT: Leach, Martin D
;; APPLICANT: Shinkets, Richard A
;; TITLE OF INVENTION: No. 6875570el Proteins and Nucleic Acids Encoding Same
;; FILE REFERENCE: 21402-124
;; CURRENT APPLICATION NUMBER: US/09/964,956
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 60/235,631
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/235,633
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/235,808
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,064
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,065
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,066
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/236,135
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: 60/237,434
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/238,321
;; PRIOR FILING DATE: 2000-10-05
;; PRIOR APPLICATION NUMBER: 60/238,399
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/238,396
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/276,667
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/294,823
;; PRIOR FILING DATE: 2001-05-31
;; PRIOR APPLICATION NUMBER: 60/304,868
;; PRIOR FILING DATE: 2001-07-12
;; NUMBER OF SEQ ID NOS: 127
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 17
;; LENGTH: 879
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-964-956-17

Query Match 7.8%; Score 145.5; DB 2; Length 879;
Best Local Similarity 23.5%; Pred. No. 1.5e-06;
Matches 68; Conservative 49; Mismatches 127; Indels 45; Gaps 10;

QY 8 GCRNGLSKYRCLDKAEANGIVLETATAGVVTSAFMTLTPILVCKVQDSNR----- 61
DB 561 GCYN-LPEDYIRWED---AWAIGPVTIACLG-----FMCTCIVITVFIKHNTPLVKAS 610

QY 62 -RKMLPTQFLFLGLVGIPLTFAFIIGLDGSTGPTFRFLFGILFSCISCLLAHAVSLT 120
DB 611 GRELC---YILLFGVSLSYCMTEFFFAKPSVICALRLRLGLGTSFAICYSALLTKTNCA 667

QY 121 KLVGRGK-----PLSLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNFSE--- 167
DB 668 RIFDGVNGAQRPKFISPSQVFI---CLGLILVQIVMVSVMLILETPGTRRYTLPEKRE 724

QY 168 ---LSAPRNEDFVLLTY--VLFMLALTFLMSSFTFCGSGFTGKRGHGAHIYLTMLLSIA 222
DB 725 TVILKCNVNDSSMLISLTYDVVLVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCI--- 780

QY 62 -RKMLPTQFLVLGLVIGLFTAFIIGLDGSGTPTRFFLFGILFSCFSLAHAVSLT 120
Db 611 GRELC---YLLFGVGLSYCMTHFFIAKPSVICALRRGLGSSFAICYSALLTKNCIA 667
QY 121 KLVGRK-----PLSLVLGLVAGFSLVQDVIAIEYIVLTWNRNINVNSE--- 167
Db 668 RIFDGVKNGAQRKPFISPSQVFI---CLGLILVQIVMVSVMLLEAPGTRRYTLAEKRE 724
QY 168 ---LSAPRRNEDFVLLLT---VLFLMALTFMLSSFTFCGSGFTGKRGHAIYLTMLLSIA 222
Db 725 TVILKCNVDSSMLISLTVDVILCTVYAFKRCPENFNKAFIGFTMYTTCI---- 780
QY 223 IWAVITLMLPDRRWDITLSSALAANGWVFLLAYVSPFEWLLTKQ 271
Db 781 IWLAFPIFYVTSSDYRVQTTMCISVLSLGGFVVLGCLFAPKVHILFQ 829

RESULT 13

US-08-486-270-6
; Sequence 6, Application US/08486270
; Patent No. 5807689
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; APPLICANT: Johnson, Edwin C.
; APPLICANT: Hees, Stephen D.
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,270
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/072,574
; FILING DATE: 04-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: FP41 9772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-270-6

Query Match 7.5%; Score 139.5; DB 1; Length 879;
Best Local Similarity 22.8%; Pred. No. 6.6e-06;
Matches 66; Conservative 52; Mismatches 126; Indels 45; Gaps 10;
QY 8 GCRNGLKSKYRLCDKABAWGIVLETATAGVVTSVAFMLTLPILVCKVQDSNR----- 61
Db 561 GCYD-LPEDIYRWD---AWAIGPVTIACLG-----FMCTCMVTVTFIKHNNTPLVKAS 610

QY 62 -RKMLPTQFLVLGLVIGLFTAFIIGLDGSGTPTRFFLFGILFSCFSLAHAVSLT 120
Db 611 GRELC---YLLFGVGLSYCMTHFFIAKPSVICALRRGLGSSFAICYSALLTKNCIA 667
QY 121 KLVGRK-----PLSLVLGLVAGFSLVQDVIAIEYIVLTWNRNINVNSE--- 167
Db 668 RIFDGVKNGAQRKPFISPSQVFI---CLGLILVQIVMVSVMLLEAPGTRRYTLAEKRE 724
QY 168 ---LSAPRRNEDFVLLLT---VLFLMALTFMLSSFTFCGSGFTGKRGHAIYLTMLLSIA 222
Db 725 TVILKCNVDSSMLISLTVDVILCTVYAFKRCPENFNKAFIGFTMYTTCI---- 780
QY 223 IWAVITLMLPDRRWDITLSSALAANGWVFLLAYVSPFEWLLTKQ 271
Db 781 IWLAFPIFYVTSSDYRVQTTMCISVLSLGGFVVLGCLFAPKVHILFQ 829

RESULT 14

US-08-367-264-6
; Sequence 6, Application US/08367264
; Patent No. 6001581
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; APPLICANT: Johnson, Edwin C.
; APPLICANT: Hees, Stephen D.
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367,264
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/072,574
; FILING DATE: 04-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: FP41 9772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-367-264-6

Query Match 7.5%; Score 139.5; DB 2; Length 879;
Best Local Similarity 22.8%; Pred. No. 6.6e-06;
Matches 66; Conservative 52; Mismatches 126; Indels 45; Gaps 10;
QY 8 GCRNGLKSKYRLCDKABAWGIVLETATAGVVTSVAFMLTLPILVCKVQDSNR----- 61
Db 561 GCYD-LPEDIYRWD---AWAIGPVTIACLG-----FMCTCMVTVTFIKHNNTPLVKAS 610

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:02:46 ; Search time 11516 Seconds
(without alignments)
11925.832 Million cell updates/sec

Title: US-10-600-816-2
Perfect score: 2456
Sequence: 1 ataacagcatgaagtgcggt.....ggcagcaaaaaaaaaa 2456

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est3: *
3: gb_est4: *
4: gb_est5: *
5: gb_est6: *
6: gb_hic: *
7: gb_est2: *
8: gb_est7: *
9: gb_est8: *
10: gb_est9: *
11: gb_gsa1: *
12: gb_gsa2: *
13: gb_gsa3: *
14: gb_gsa4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1830	74.5	1830	6	CR597125	CR597125 full-length
2	1037.6	42.2	1740	6	CR620367	CR620367 full-length
3	965.8	39.3	1022	1	AL556103	AL556103 AL556103
4	935	38.1	1004	1	AL556722	AL556722 AL556722
5	929.4	37.8	1066	1	AL544370	AL544370 AL544370
6	916.8	37.3	1106	1	AL578517	AL578517 AL578517
7	867.4	35.3	902	3	BU552514	BU552514 AGENCOURT
8	864	35.2	1040	3	BM921746	BM921746 AGENCOURT
9	855.2	34.8	983	3	BQ958916	BQ958916 AGENCOURT
10	851.6	34.7	988	3	BM921595	BM921595 AGENCOURT
11	851.2	34.7	862	1	AL571607	AL571607 AL571607
12	843.8	34.4	899	3	BQ936508	BQ936508 AGENCOURT
13	834.8	34.0	897	2	BI820695	BI820695 603034482
14	829.8	33.8	916	5	CF780868	CF780868 AGENCOURT
15	820	33.4	839	1	AL546138	AL546138 AL546138
16	819.8	33.4	918	4	BX372160	BX372160 BX372160
17	817.8	33.3	866	2	BI765773	BI765773 603046537
18	815.4	33.2	1023	3	BQ067433	BQ067433 AGENCOURT
19	808.2	32.9	885	3	BQ430279	BQ430279 AGENCOURT

20	803.4	32.7	899	3	BQ958341	BQ958341 AGENCOURT
21	796	32.4	796	2	BG830281	BG830281 602764675
22	789.8	32.2	812	2	BI824631	BI824631 603033572
23	786	32.0	945	2	BG752462	BG752462 602730954
24	781.2	31.8	865	4	CA454746	CA454746 AGENCOURT
25	779.2	31.7	940	3	BQ685729	BQ685729 AGENCOURT
26	777.4	31.7	933	2	BG823401	BG823401 602726823
27	774.8	31.5	804	2	BM046252	BM046252 603626039
28	771.8	31.4	920	3	BQ920279	BQ920279 AGENCOURT
29	770.4	31.4	895	3	BUI50237	BUI50237 AGENCOURT
30	769.8	31.3	1091	3	BQ922898	BQ922898 AGENCOURT
31	766	31.2	1137	2	BM545793	BM545793 AGENCOURT
32	758.4	30.9	773	1	AL709960	AL709960 DKFZp686D
33	755.8	30.8	962	4	BX370558	BX370558 BX370558
34	752.2	30.6	758	2	BM049765	BM049765 603624288
35	749.4	30.5	795	2	BM049939	BM049939 603624430
36	743.2	30.3	908	2	BG824529	BG824529 602728414
37	723.8	29.5	923	3	BU543952	BU543952 AGENCOURT
38	723.4	29.4	737	8	CN348542	CN348542 170006003
39	722	29.4	880	3	BQ944973	BQ944973 AGENCOURT
40	710.2	28.9	810	2	BI767478	BI767478 603057344
41	708	28.8	781	9	DA647581	DA647581 DA647581
42	707.2	28.8	717	9	DA647448	DA647448 DA647448
43	697	28.4	709	3	BM784295	BM784295 K-EST0062
44	696.8	28.4	788	2	BG118532	BG118532 602348467
45	683.6	27.8	702	2	BG251131	BG251131 602364932

ALIGNMENTS

RESULT 1	CR597125	1830 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CSODK011YF1.7 of Homo sapiens (human).				
ACCESSION	CR597125				
VERSION	CR597125.1 GI:50477932				
KEYWORDS	HTC; CNSLT_cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1830)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 1830)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]				
COMMENT	- Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..1830				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSODK011YF1.7"				
	/tissue_type="Hela cells Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	74.5%	Score	1830;	DB 6;	Length 1830;

		Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps	0;
		Matches	1830;	Conservative	0;					
Qy	154	TCCAAGGTCTCCCCCACGACTGAGGAGCTCGCCTGTGCCTCTTTGGCGCGGGAAGCAG	213							
D _b	1	TCCNAGGTCTCCCCACGACTGAGGAGCTGCCTGTGCCCTCTTTGGCGCGGGAAGCAG	60							
Qy	214	CACCAAGTTCACGGCCAAACGCTTTGGCACATAGGTCAGGAATGGCTACAACAGTCCCTGA	273							
D _b	61	CACCAAGTTCAGGSCCAAACGCTTTGGCACATAGGTCAGGAATGGCTACAACAGTCCCTGA	120							
Qy	274	TGGTTGCCGCAATGGCCTGAAAATCCAAGTAATAAGCTATTGTGTGAATAGGCTGAAGCTTG	333							
D _b	121	TGGTTGCCGCAATGGCCTGAAAATCCAAGTAATAAGCTATTGTGTGAATAGGCTGAAGCTTG	180							
Qy	334	GGGCATCGTCTAGAAACGCTGGCCACAGCCGGGTTGTACCTCGGTGGCCTTCATGCT	393							
D _b	181	GGGCATCGTCTAGAAACGCTGGCCACAGCCGGGTTGTACCTCGGTGGCCTTCATGCT	240							
Qy	394	CACCTCTCCGATCCTCGTCTGCAAGTTCAGGACTCCAACAGGCGAAAAATGCTGCTAC	453							
D _b	241	CACCTCTCCGATCCTCGTCTGCAAGTTCAGGACTCCAACAGGCGAAAAATGCTGCTAC	300							
Qy	454	TCAGTTTTCTTCTCTCTGGGTGTGTTGGGCAATTTTGGCCTCACCTTCGCTTCATCAT	513							
D _b	301	TCAGTTTTCTTCTCTCTGGGTGTGTTGGGCAATTTTGGCCTCACCTTCGCTTCATCAT	360							
Qy	514	CGGACTCGAAGGGAGACAGGGCCACACGCTCTTCCTCTTTGGGATCCTCTTTTCCAT	573							
D _b	361	CGGACTCGAAGGGAGACAGGGCCACACGCTCTTCCTCTTTGGGATCCTCTTTTCCAT	420							
Qy	574	CTGCTTCTCTGCTGCTGGCTCATGCTGTCAAGTCTACCAAGCTCGTCGGGGGAGGAA	633							
D _b	421	CTGCTTCTCTGCTGCTGGCTCATGCTGTCAAGTCTACCAAGCTCGTCGGGGGAGGAA	480							
Qy	634	GCOCCTTTCCTGTPTGTGTGATTCGTGGGTCTGGCGGTGGGCTTCAGCCTAGTCAGAGATG	693							
D _b	481	GCOCCTTTCCTGTPTGTGTGATTCGTGGGTCTGGCGGTGGGCTTCAGCCTAGTCAGAGATG	540							
Qy	694	TATCGCTTATGAAATATATTGCTCTGACCATGATAGGACCACGCTCAATGCTCTTTCTGA	753							
D _b	541	TATCGCTTATGAAATATATTGCTCTGACCATGATAGGACCACGCTCAATGCTCTTTCTGA	600							
Qy	754	GCTTTCGGCTCCTCGTCGCAATGAAGACTTTGTCTCTGCTCACCTACCTACGCTCTCTCTT	813							
D _b	601	GCTTTCGGCTCCTCGTCGCAATGAAGACTTTGTCTCTGCTCACCTACCTACGCTCTCTT	660							
Qy	814	GATGGCGCTGACCTTCTCTCATGTCTCCTTCAOCTTGTGGTTCCTTTCACGGGCTGGAA	873							
D _b	661	GATGGCGCTGACCTTCTCTCATGTCTCCTTCAOCTTGTGGTTCCTTTCACGGGCTGGAA	720							
Qy	874	GAGACATGGGGCCACATCTACCTCAGATGCTCCTCTCCATTTGCCATCTGGGTGGCCTG	933							
D _b	721	GAGACATGGGGCCACATCTACCTCAGATGCTCCTCTCCATTTGCCATCTGGGTGGCCTG	780							
Qy	934	GATCACCTCTCATGTCTCTGACTTTTGACCGCAGGTGGATGACACCATCCTCAGCTC	993							
D _b	781	GATCACCTCTCATGTCTCTGACTTTTGACCGCAGGTGGATGACACCATCCTCAGCTC	840							
Qy	994	CGCCTTGGCTGCCAATGGCTGGGTGTTTCTGTGGCTTATGTTAGTCCCGAGTTTGGCT	1053							
D _b	841	CGCCTTGGCTGCCAATGGCTGGGTGTTTCTGTGGCTTATGTTAGTCCCGAGTTTGGCT	900							
Qy	1054	GCTCACAAGCAACGAAACCCCATGGATATATCTGTTGAGGATGCTTTCTGTAACCTCA	1113							
D _b	901	GCTCACAAGCAACGAAACCCCATGGATATATCTGTTGAGGATGCTTTCTGTAACCTCA	960							
Qy	1114	ACTCGTGAAGAAGAGCTATGTTGTGGAGAACAGAGGCTACTCTCAAGAGGAAATCACTCA	1173							
D _b	961	ACTCGTGAAGAAGAGCTATGTTGTGGAGAACAGAGGCTACTCTCAAGAGGAAATCACTCA	1020							
Qy	1174	AGSTTTTGAAGAGACAGGGGACAGCTCTATGCCCCCTATTCCACATTTTTCAGTGCAC	1233							

Db	1021	AGTTTTGAAGAGACAGGGGACAGCTCTATGCGCCCTATTCCACACATTTTCAGCTGCA	1080
Qy	1234	GAACCAAGCCTCCCAAAAGGAAATTTCTCCATCCACGGGGCCACGCTTTGGCCGAGCCCTTA	1293
Db	1081	GAACCAAGCCTCCCAAAAGGAAATTTCTCCATCCACGGGGCCACGCTTTGGCCGAGCCCTTA	1140
Qy	1294	CAAAGACTATGAAGTAAAGAAAAGAGGCGACTAACTCTGTCTCTGAAGAGTGGGACAAATG	1353
Db	1141	CAAAGACTATGAAGTAAAGAAAAGAGGCGACTAACTCTGTCTCTGAAGAGTGGGACAAATG	1200
Qy	1354	CAGCCGGCGGCGAGATCTACGGGAGGCTCAAGAGGATGTGGCGGAAATCTTGAGTCTTCT	1413
Db	1201	CAGCCGGCGGCGAGATCTACGGGAGGCTCAAGAGGATGTGGCGGAAATCTTGAGTCTTCT	1260
Qy	1414	GAGAAACTCTCAAGACACATAGGGGAAACAGTTTTCCTCCCTCCAGCCTCAACACAAT	1473
Db	1261	GAGAAACTCTCAAGACACATAGGGGAAACAGTTTTCCTCCCTCCAGCCTCAACACAAT	1320
Qy	1474	TCTTCCATGTGGGGCTGATGTGGGCTAGTAAAGACTCCAGTTCTTAGAGGCGCTGTAGTA	1533
Db	1321	TCTTCCATGTGGGGCTGATGTGGGCTAGTAAAGACTCCAGTTCTTAGAGGCGCTGTAGTA	1380
Qy	1534	TTTTTTTTTTTTTGTCTCATCTCTTTTGATCTCTTTTAACTGGGAGTCTCAGGCAACTC	1593
Db	1381	TTTTTTTTTTTTTGTCTCATCTCTTTTGATCTCTTTTAACTGGGAGTCTCAGGCAACTC	1440
Qy	1594	AAGTTTAGACCCCTTACTCTCTTTTGTCTTTTGTAAACAGGATCTTGCTCTGTCAACCA	1653
Db	1441	AAGTTTAGACCCCTTACTCTCTTTTGTCTTTTGTAAACAGGATCTTGCTCTGTCAACCA	1500
Qy	1654	GGCTTGAGTGCAGTGTGCGATCACAGCCAGTGCAGCCTCGACACCTGTGTCTCAAGCA	1713
Db	1501	GGCTTGAGTGCAGTGTGCGATCACAGCCAGTGCAGCCTCGACACCTGTGTCTCAAGCA	1560
Qy	1714	ATCCTCCACTCTCCATCTCCCAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCCAGC	1773
Db	1561	ATCCTCCACTCTCCATCTCCCAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCCAGC	1620
Qy	1774	CTAGGCCCTTAAATCTTGCTGTATTTTCCATGGAATAAAGTCTGGTCACTCTCAGCTCAC	1833
Db	1621	CTAGGCCCTTAAATCTTGCTGTATTTTCCATGGAATAAAGTCTGGTCACTCTCAGCTCAC	1680
Qy	1834	GCTGGCTCACACAGCTCTAGGGGCGCTCTCTCTTAATCAAGTGGGTTTGTGAGGCTC	1893
Db	1681	GCTGGCTCACACAGCTCTAGGGGCGCTCTCTCTTAATCAAGTGGGTTTGTGAGGCTC	1740
Qy	1894	TGTGGCCAGAGCAGACTCGATATCTGAGCAAAATAGCAAAAGCCTCTCTCAGGCCAC	1953
Db	1741	TGTGGCCAGAGCAGACTCGATATCTGAGCAAAATAGCAAAAGCCTCTCTCAGGCCAC	1800
Qy	1954	TGGCCTGAATCTACACTGGAAGCAACTTG	1983
Db	1801	TGGCCTGAATCTACACTGGAAGCAACTTG	1930
RESULT 2			
CR620367		1740 bp mRNA linear HTC 21-JUL-2004	
LOCUS		full-length cDNA clone CSODI018P20 of Placenta Cot 25-normalized	
DEFINITION		of Homo sapiens (human).	
ACCESSION		CR620367	
VERSION		CR620367.1 GI:50501174	
KEYWORDS		HTC; CNSLT cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		1 (bases 1 to 1740)	
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE		Full-length cDNA libraries and normalization	
JOURNAL		Unpublished	
REMARK		Contact : Feng Liang Email : fliang@lifetech.com URL :	

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1740)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1018YP20"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 42.2%; Score 1037.6; DB 6; Length 1740;
Best Local Similarity 99.1%; Pred. No. 1.5e-209;
Matches 1043; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 141 GCGGGATAGCTGTCCAAAGGTCTCCCCAGCACTGAGGAGCTCGCTGTGCCCTCTTGC 200
DB 1 GCGGGATAGCTGTCCAAAGGTCTCCCCAGCACTGAGGAGCTCGCTGTGCCCTCTTGC 60
QY 201 GCGCGGAGAGCAGCACAAGTTACAGGCCAAGCCCTTGGCACTAGGCTCCAGAAAGGCTA 260
DB 61 GCGCGGAGAGCAGCACAAGTTACAGGCCAAGCCCTTGGCACTAGGCTCCAGAAAGGCTA 120
QY 261 CAACAGTCCCTGATGTTGCGCAATGGCTCGAAATCCAAAGTACTACAGACTTTGTGATA 320
DB 121 CAACAGTCCCTGATGTTGCGCAATGGCTCGAAATCCAAAGTACTACAGACTTTGTGATA 180
QY 321 AGCTGAGCTTGGGCGATCGCTAGAAAAGGTGGCCACAGCGGGGTGTGACCTCGG 380
DB 181 AGCTGAGCTTGGGCGATCGCTAGAAAAGGTGGCCACAGCGGGGTGTGACCTCGG 240
QY 381 TGGCTTCTATGCTCACTCTCCGATCTCTGCTGCAAGGTGAGGACTCCACAGCGGAA 440
DB 241 TGGCTTCTATGCTCACTCTCCGATCTCTGCTGCAAGGTGAGGACTCCACAGCGGAA 300
QY 441 AAATGCTGCTACTCAGTTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
DB 301 AAATGCTGCTACTCAGTTTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 501 TGGCTTCTATGCTGAGAGGAGCAGAGGCGCCACAGCTTCTCTCTCTCTCTCTCTCT 560
DB 361 TGGCTTCTATGCTGAGAGGAGCAGAGGCGCCACAGCTTCTCTCTCTCTCTCTCTCT 420
QY 561 TCGCTTTTCCATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620
DB 421 TCGCTTTTCCATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 621 TCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 680
DB 481 TCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 681 TAGTCAGAGATGTTATCGCTATTTGAATATATTTGCTCTGACCAATGAATAGACCAAGTCA 740
DB 541 TAGTCAGAGATGTTATCGCTATTTGAATATATTTGCTCTGACCAATGAATAGACCAAGTCA 600
QY 741 ATGCTCTTCTGAGCTTTCGGCTTCCTCGCAAGCAAGCTTTCCTCTCTCTCTCTCTCT 800
DB 601 ATGCTCTTCTGAGCTTTCGGCTTCCTCGCAAGCAAGCTTTCCTCTCTCTCTCTCTCT 660
QY 801 AGCTCTCTCTTGTAGGGCGCTGACCTCTCATGTCCTCTCTCTCTCTCTCTCTCTCTCT 860

DB 661 ACCTCTCTCTTCTTGATGGGCTGACCTTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 861 TCACGGCTGGAAGAGACATGGGGCCCACTTACTCTACGATGCTCTCTCTCTCTCTCTCTCT 920
DB 721 TCACGGCTGGAAGAGACATGGGGCCCACTTACTCTACGATGCTCTCTCTCTCTCTCTCT 780
QY 921 TCTGGGTGGCTTGGATCACCTGCTCATGCTTCTCTGACTTTGACCCGAGGTGGGATGACA 980
DB 781 TCTGGGTGGCTTGGATCACCTGCTCATGCTTCTCTGACTTTGACCCGAGGTGGGATGACA 840
QY 981 CCATCTCAGCTCCGCTTGGCTTGGCAATGGCTGGGTGCTCTCTCTCTCTCTCTCTCTCTCT 1040
DB 841 CCATCTCAGCTCCGCTTGGCTTGGCAATGGCTGGGTGCTCTCTCTCTCTCTCTCTCTCTCT 900
QY 1041 CCGAGTTTGGCTGCTCACAAGACGAAACCCCATGCTTATCTCTCTCTCTCTCTCTCTCTCT 1100
DB 901 CCGAGTTTGGCTGCTCACAAGACGAAACCCCATGCTTATCTCTCTCTCTCTCTCTCTCTCT 960
QY 1101 TCTGTAACCTCAACTCTGTAAGAGAGCTATGGTGTGGAGAACAGAGCTTCTCTCAAG 1160
DB 961 TCTGTAACCTCAACTCTGTAAGAGAGCTATGGTGTGGAGAACAGAGCTTCTCTCAAG 1020
QY 1161 AGGAATCACTCAAGGTTTGAAGACAGGG 1192
DB 1021 AGGAATCACTCAAGGTCAGATGCGAGCTGG 1052
RESULT 3
AL556103 1022 bp mRNA linear EST 30-MAR-2004
LOCUS AL556103 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK011YF17 5-PRIME, mRNA sequence.
ACCESSION AL556103
VERSION AL556103.3 GI:45860822
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31277907.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6601.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK011CC09QP1&c=6601.f.
FEATURES
Location/Qualifiers
1..1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YF17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was
normalized." sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 39.3%; Score 965.8; DB 1; Length 1022;

Best Local Similarity 98.7%; Pred. No. 2.6e-194; Matches 1000; Conservative 4; Mismatches 6; Indels 3; Gaps 3;

AL556722
LOCUS AL556722 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens EST 02-APR-2000
DEFINITION CDNA clone CS0DK005Y006 5-PRIME, mRNA sequence.

ACCESSION AL556722 GI:46182140
VERSION AL556722.3
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31278523.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6601.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?cs=CS0DK005BH03QP1&c=6601.f>.

FEATURES
source
1..1004
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK005Y006"
/cell_line="HELA"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the PCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 38.1%; Score 935; DB 1; Length 1004;
Best Local Similarity 97.3%; Pred. No. 9.2e-188;
Matches 959; Conservative 3; Mismatches 23; Indels 1; Gaps 1;

QY 152 TGTCGAAGGCTCCCCCAGCACTGAGAGCTCCCTGTGCCTCTTTGCCGGCGGGAAGC 211
DB 1 TGTCCAAGGCTCCCCCAGCACTGAGAGCTCCCTGTGCCTCTTTGCCGGCGGGAAGC 60

QY 212 AGCACCAAGTTACGGCCCAACGCTTGCGACTAGGCTCCAGATGGCTACAACAGTCCCT 271
DB 61 AGCACCAAGTTACGGCCCAACGCTTGCGACTAGGCTCCAGATGGCTACAACAGTCCCT 120

QY 272 GATGGTTGCGCAATGGCCCTGAAATCCAAGTACTACAGACTTTGTGATAAGGCTGAAGCT 331
DB 121 GATGGTTGCGCAATGGCCCTGAAATCCAAGTACTACAGACTTTGTGATAAGGCTGAAGCT 180

QY 332 TGGGGCATCGTCTTAGAAAACGGTGGCCACAGCCGGGGTTGTGACCTCGGTGGGCTTCATG 391
DB 181 TGGGGCATCGTCTTAGAAAACGGTGGCCACAGCCGGGGTTGTGACCTCGGTGGGCTTCATG 240

QY 392 CTCACTCTCCGATCTCTAGAAAACGGTGGCCACAGCCGGGGTTGTGACCTCGGTGGGCTTCATG 451
DB 241 CTCACTCTCCGATCTCTAGAAAACGGTGGCCACAGCCGGGGTTGTGACCTCGGTGGGCTTCATG 300

QY 452 ACTCAGTTTTCTCTCTCTGGGTGTGTGGGCATCTTTGGCTTCACCTTCGGCTTCATC 511
DB 301 ACTCAGTTTTCTCTCTCTGGGTGTGTGGGCATCTTTGGCTTCACCTTCGGCTTCATC 360

QY 512 ATCGGACTGGACGGGAGCACACAGGCCCCACACGGCTTCTTCTCTTTGGGATCTCTTTTTC 571

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2774 row: j column: 17
 High quality sequence stop: 716.
 Location/Qualifiers
 1. .902
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6576305"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 35.3%; Score 867.4; DB 3; Length 902;
 Best Local Similarity 98.4%; Pred. No. 2.1e-173;
 Matches 887; Conservative 0; Mismatches 11; Indels 3; Gaps 1;

QY 164 CCCCCAGCTAGGAGCTGCCTGCTGCCCTCTTCGCGCGGGAAGCAGCACCAAGTTC 223
 DB 1 CCCCCAGCTAGGAGCTGCCTGCTGCCCTCTTCGCGCGGGAAGCAGCACCAAGTTC 60

QY 224 ACGGCCAAGCCCTGGCAGCTAGGGTCCAGATGGCTACACAGTCCCTGATGGTTCGCG 283
 DB 61 ACGGCCAAGCCCTGGCAGCTAGGGTCCAGATGGCTACACAGTCCCTGATGGTTCGCG 120

QY 284 AATGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGCTGAAGCTTGGGCGATCGTC 343
 DB 121 AATGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGCTGAAGCTTGGGCGATCGTC 180

QY 344 CTAGAAACGGTGGCCACAGCGGGGTGTGACCTCGGTGGCCCTTCATGCTCACTCTCCCG 403
 DB 181 CTAGAAACGGTGGCCACAGCGGGGTGTGACCTCGGTGGCCCTTCATGCTCACTCTCCCG 240

QY 404 ATCTCTGCTGCAAGGTGAGGACTCCACAGCGGAAAATGCTGCTACTAGTTTCTC 463
 DB 241 ATCTCTGCTGCAAGGTGAGGACTCCACAGCGGAAAATGCTGCTACTAGTTTCTC 300

QY 464 TTCTCTCTGGGTGTGGGATCTTTGGCTCACCCTTCGCTTCATCATCGAGCTGGAC 523
 DB 301 TTCTCTCTGGGTGTGGGATCTTTGGCTCACCCTTCGCTTCATCATCGAGCTGGAC 360

QY 524 GGGAGCACAGGGCCCCACACGCTCTTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCC 583
 DB 361 GGGAGCACAGGGCCCCACACGCTCTTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCC 420

QY 584 TGCTCTGCTGCTATGCTGTGCTGCTGACCAAGCTGTCGGGGGAGGAGCCCTTTCC 643
 DB 421 TGCTCTGCTGCTATGCTGTGCTGCTGACCAAGCTGTCGGGGGAGGAGCCCTTTCC 480

QY 644 CTCTTGTGATCTGGGTCTGGCGGTGGGCTTCAGCTAGTCAGAGATGTTATCGCTATT 703
 DB 481 CTCTTGTGATCTGGGTCTGGCGGTGGGCTTCAGCTAGTCAGAGATGTTATCGCTATT 540

QY 704 GAATATATTGCTCTGACCAATGAATAGGACCAACGTCATATGCTCTTTTCTGAGCTTTCCGCT 763
 DB 541 GAATATATTGCTCTGACCAATGAATAGGACCAACGTCATATGCTCTTTTCTGAGCTTTCCGCT 600

QY 764 CCTCGTGGCAATGAAGACCTTTCTCTCTGCTCACTACGTCCTCTTCTTTGATGGCGCTG 823
 DB 601 CCTCGTGGCAATGAAGACCTTTCTCTCTGCTCACTACGTCCTCTTCTTTGATGGCGCTG 660

QY 824 ACCTTCCTCATGCTCCTCTCTCTTCCACCTTCTGTGGTTCCTTCACGGGCTGGAAGACATGGG 883
 DB 661 ACCTTCCTCATGCTCCTCTCTCTTCCACCTTCTGTGGTTCCTTCACGGGCTGGAAGACATGGG 720

QY 884 GCCACATCTACCTCACGATGCTCTCTCCATTCGATGCCATCTGGGTGGCTGGATCACCCCTG 943
 DB 721 GCCACATCTACCTCACGATGCTCTCTCCATTCGATGCCATCTGGGTGGCTGGATCACCCCTG 780

QY 944 CTCATGCTCTCTGACTTTGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTTGGCT 1003
 DB 781 CTCATGCTCTCTGACTTTGACCGCAGGTGGGATGACACCATCTCAGCTCCGCTTTGGCT 840

QY 1004 GCCAATGGCTGGGTGTTCTCTGTGGCTTATGTT--AGTCCCGAGTTTTGGCTGCTCACA 1060
 DB 841 GCCAATGGCTGGGTGTTCTCTGTGGCTATGTTAGTCCCAAGTTTGGCTGCTCACA 900

QY 1061 A 1061
 DB 901 A 901

RESULT 8
 BM921746
 LOCUS
 DEFINITION
 5', mRNA sequence.
 ACCESSION
 BM921746
 VERSION
 BM921746.1 GI:19372125
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1040)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12788 row: o column: 13
 High quality sequence stop: 623.
 Location/Qualifiers
 1. .1040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5753316"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

FEATURES
 source

ORIGIN

Query Match		35.2%;	Score 864;	DB 3;	Length 1040;
Best Local Similarity		95.8%;	Pred. No. 1.1e-172;		
Matches 908;		Conservative 0;	Mismatches 37;	Indels 3;	Gaps 2;
QY	161	TCCTCCCCAGCACTGAGAGCTCGCTGCTGCCCTTTCGCGCGGGAAGCAGCACCAAG	220		
DB	1	TCCTCCCCAGCACTGAGAGCTCGCTGCTGCCCTTTCGCGCGGGAAGCAGCACCAAG	60		
QY	221	TTTCAGGGCCAAACGCTTGGCACTAGGCTCCAGATGGCTTACACAGTCCCTGATGTTGC	280		
DB	61	TTTCAGGGCCAAACGCTTGGCACTAGGCTCCAGATGGCTTACACAGTCCCTGATGTTTC	120		
QY	281	CGCAATGGCTGAATCCAAAGTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATC	340		
DB	121	CGCAATGGCTGAATCCAAAGTACTACAGACTTTGTGATAAGCTGAAGCTTGGGGCATC	180		
QY	341	GTCTAGAAACGGTGGCCACAGCGGGTGTGACTCGTGGCTTCAATGCTCACTCTC	400		
DB	181	GTCTAGAAACGGTGGCCACAGCGGGTGTGACTCGTGGCTTCAATGCTCACTCTC	240		
QY	401	CGGATCTCTGCTGCAAGTGCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTT	460		
DB	241	CGGATCTCTGCTGCAAGTGCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTT	300		
QY	461	CTCTTCTCTCGGTGTGTGGGCATCTTTGGCTCACCCTTCGCTTCAATCGGACTG	520		
DB	301	CTCTTCTCTCGGTGTGTGGGCATCTTTGGCTCACCCTTCGCTTCAATCGGACTG	360		
QY	521	GACGGGACACAGGCCCCACAGCTCTTCTTCTTGGGATCCTCTTTTCAATCGCTTC	580		
DB	361	GACGGGACACAGGCCCCACAGCTCTTCTTCTTGGGATCCTCTTTTCAATCGCTTC	420		
QY	581	TCCTGCTCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	640		
DB	421	TCCTGCTCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	480		
QY	641	TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	700		
DB	481	TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540		
QY	701	ATTGAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	760		
DB	541	ATTGAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600		
QY	761	GCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	820		
DB	601	GCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660		
QY	821	CTGACCTTCT	880		
DB	661	CTGACCTTCT	720		
QY	881	GGGGCCACATCTACTACAGATGCT	940		
DB	721	GGGGCCACATCTACTACAGATGCT	780		
QY	941	CTGCTCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1000		
DB	781	CTGCTCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840		
QY	1001	GCTGCAATGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1057		
DB	841	GCTGCAATGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900		
QY	1058	ACAAAGCAACGAAACCCCGGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1105		
DB	901	CCAAAGCAACGAAACCCCGGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	948		

RESULT 9
BQ958916
LOCUS

BQ958916 983 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT_10029407 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481982
5', mRNA sequence.
ACCESSION BQ958916
VERSION BQ958916.1 GI:22374394
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 983)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/Drp
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2664 row: d column: 15
High quality sequence stop: 615.
Location/Qualifiers
1. 983
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481982"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match		34.8%;	Score 855.2;	DB 3;	Length 983;
Best Local Similarity		96.4%;	Pred. No. 8.3e-171;		
Matches 906;		Conservative 0;	Mismatches 30;	Indels 4;	Gaps 3;
QY	169	AGCACTGAGGAGCTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGCACCAAGTTTCAACGCG	228		
DB	1	AGCACTGAGGAGCTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGCACCAAGTTTCAACGCG	60		
QY	229	CAACGCTTGGCACTAGGGTCCAGAAATGGCTACAAACAGTCCCTGATGGTTGCGCAATGG	288		
DB	61	CAACGCTTGGCACTAGGGTCCAGAAATGGCTACAAACAGTCCCTGATGGTTGCGCAATGG	120		
QY	289	CCTGAAATCCAAAGTACTACAGCTTTGTGATAAGCTGGAAGCTTGGGGCATCGTCTCTAGA	348		
DB	121	CCTGAAATCCAAAGTACTACAGCTTTGTGATAAGCTGGAAGCTTGGGGCATCGTCTCTAGA	180		
QY	349	AACGCTGGCCACAGCGGGGTGTGACCTCGTGGGCTTTCATGCTCACTCTCCCGATCCT	408		
DB	181	AACGCTGGCCACAGCGGGGTGTGACCTCGTGGGCTTTCATGCTCACTCTCCCGATCCT	240		
QY	409	CGTCTGCAAGTGCAGGACTCCAAACAGCGGAAAAATGCTGCTACTCAGTTTCTCTTCT	468		
DB	241	CGTCTGCAAGTGCAGGACTCCAAACAGCGGAAAAATGCTGCTACTCAGTTTCTCTTCT	300		
QY	469	CCTGGGTGTGTGGGCATCTTTGGCTCACTTCCCTTTCATCGACTCGGACGGGAG	528		
DB	301	CCTGGGTGTGTGGGCATCTTTGGCTCACTTCCCTTTCATCGACTCGGACGGGAG	360		
QY	529	CACAGGGCCACACGCTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTCTGCT	588		


```
Db      780 CTGCTCATGCTCTCTGACNTGACCGCAGGTGGGATGACACCATCTCTAGCTCCGCCCTTG 839
Qy      1001 GCTGCAATGGCTGGG-TGTTCTCTGTGGCTTATGTATGTCCTCCGAG---TTTGGCTGCT 1056
Db      840 GCTGCAATGGCTGGGTTGCTCTGCTGCTTATGTTTANTCCCGAGTTTTTGGCTGCT 899
Qy      1057 CACAAA-GCAAGAACCCCATGGATTAT 1084
Db      900 CACAAAGGCAACGAAACCCCTGGGATTT 928

RESULT 11
AL571607/c
LOCUS      AL571607 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1025YF10 3-PRIME, mRNA sequence.
ACCESSION AL571607
VERSION    AL571607.1 GI:12929072
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 862)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT    Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
            Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            6601.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?8=CS0D1025D05NP1&c=6601.f.

FEATURES             source
            Location/Qualifiers
            1..862
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0D1025YF10"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and cloned into the Not I and EcoR V
            sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      34.7%; Score 851.2; DB 1; Length 862;
Beat Local Similarity 98.6%; Pred. No. 5.8e-170;
Matches 850; Conserved 7; Mismatches 5; Indels 0; Gaps 0;

Qy      1532 TATTTTATTTTGTCTCATCTCTTGATCTCTTTTAAAGTGGAGTCTCAGGCAAC 1591
Db      862 TATTTTATTTTGTCTCATCTCTTGATCTCTTTTAAAGTGGAGTCTCAGGCAAC 803
Qy      1592 TCAAGTTAGACCCCTACTCTTTTGTGTTTTTGAACAGGATCTTGTCTGTCCACC 1651
Db      802 TCAAGTTAGACCCCTACTCTTTTGTGTTTTTGAACAGGATCTTGTCTGTCCACC 743
Qy      1652 CAGGCTTGAGTCAGTGGTGCATCACGCCAGTGCAGCTCGACCACTGTGCTCAAG 1711
Db      742 CAGGCTTGAGTCAGTGGTGCATCACGCCAGTGCAGCTCGACCACTGTGCTCAAG 683
Qx      1712 CAATCTCCCATCTCCATCTCCCAAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCCA 1771
Db      682 CAATCTCCCATCTCCATCTCCCAAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCCA 623
```

```
Qy      1772 GCCTAGGCCCTTAATCTTTGCTGTGTTATTTTCCATGACTAAAGGTCTGGTCACTCAGCTC 1831
Db      622 GCCTAGGCCCTTAATCTTTGCTGTGTTATTTTCCATGACTAAAGGTCTGGTCACTCAGCTC 563
Qy      1832 AGCTGGCTCACACAGCTCTAGGGGCTGCTCTCTAACTCAGTGGGTTTGTGAGGC 1891
Db      562 AGCTGGCTCACACAGCTCTAGGGGCTGCTCTCTAACTCAGTGGGTTTGTGAGGC 503
Qy      1892 TCTGTGGCCAGAGCAGACCTGTCATATCTGAGCAAAAATAGCAAAAGCCCTCTCTCAGCCC 1951
Db      502 TCTGTGGCCAGAGCAGACCTGTCATATCTGAGCAAAAATAGCAAAAGCCCTCTCTCAGCCC 443
Qy      1952 ACTGCGCTGAATCTAATCTGGAAGCAACTTGTGGGACCCCGCTCCCAACCTTCTT 2011
Db      442 ACTGCGCTGAATCTAATCTGGAAGCAACTTGTGGGACCCCGCTCCCAACCTTCTT 383
Qy      2012 GCCTGGGTAGGAGGCTTAAAGATCACCCCTAAATTTACTCATCTCTAGTCTGCCTCA 2071
Db      382 GCCTGGGTAGGAGGCTTAAAGATCACCCCTAAATTTACTCATCTCTAGTCTGCCTCA 323
Qy      2072 CATTTGGGCTCAGCAGCTCCCGCAGCAACCAATTCACAGGTCAACCCCTCTCTTTTGCACGTG 2131
Db      322 CATTTGGGCTCAGCAGCTCCCGCAGCAACCAATTCACAGGTCAACCCCTCTCTTTTGCACGTG 263
Qy      2132 TCCCAAACTTGTCTCTAATTTCCGAGATCTAATCTCCCGCTAGGCTCTGCCAGGAAATCT 2191
Db      262 TCCCAAACTTGTCTCTAATTTCCGAGATCTAATCTCCCGCTAGGCTCTGCCAGGAAATCT 203
Qy      2192 TTCAGACCTCACTAGCACAAGCCCGTGTCTCTTGTGAGGAGAAATTTGTAGATCATCTCT 2251
Db      202 TTCAGACCTCACTAGCACAAGCCCGTGTCTCTTGTGAGGAGAAATTTGTAGATCATCTCT 143
Qy      2252 CACTTCAAATTCCTGGGCTGATCTCTCTCATCTTGCACCCCAACCTCTGTAAATAGA 2311
Db      142 CACTTCAAATTCCTGGGCTGATCTCTCTCATCTTGCACCCCAACCTCTGTAAATAGA 83
Qy      2312 TTTACCGGATTTACGGCTGCATCTGTAAGTGGGATGCTCTCTAATGAGGAGGTGTC 2371
Db      82 TTTACCGGATTTACGGCTGCATCTGTAAGTGGGATGCTCTCTAATGAGGAGGTGTC 23
Qy      2372 ATTGTATAATAAGTTATTCAACC 2393
Db      22 ATTGTATAATAAGTTATTOMCC 1

RESULT 12
BQ936508
LOCUS      BQ936508
DEFINITION AGENCOURT_10058421 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6480086
            5', mRNA sequence.
ACCESSION BQ936508
VERSION    BQ936508.1 GI:22351891
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 899)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/Drp
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2659 row: e column: 15
```

FEATURES	High quality sequence stop: 687.		Db	721	CACTACCTCAGATGCTCTCTCCATTGGCATGGGTGGGCTGGATCACCTGCTCAT		780
	Location/Qualifiers						
source	1..899		QY	949	GCTTCTGACTTTGACCGCAGGTGGATGACACCATCTCTAGCTCCGCTTGCTGGCAA		1008
	/organism="Homo sapiens"						
	/mol_type="mRNA"		Db	781	GCTTCTGACTTTGACCGCAGGTGGATGACACCATCTCTAGCTCCGCTTGCTGGCAA		840
	/db_xref="taxon:9606"						
	/clone="IMAGE:648086"		QY	1009	TGGCTGGG		1016
	/issue_type="carcinoma, cell line"						
	/lab_host="DH10B (phage-resistant)"		Db	841	TGGCTGGG		848
	/clone_lib="NIH_MGC_40"						
	/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		RESULT 13	BI820695	897 bp mRNA linear EST 04-OCT-2001		603034482F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175501 5', mRNA sequence.
	ACCESSION		LOCUS	DEFINITION	EST.		BI820695.1 GI:15932245
	VERSION						
	KEYWORDS		SOURCE	ORGANISM	Homo sapiens (human)		Homo sapiens
	REFERENCE		AUTHORS	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		Unpublished (1999)
	JOURNAL						
	COMMENT		FEATURES	source	Contact: Robert Strausberg, Ph.D.		Email: cgapbs-remail.nih.gov
					Tissue Procurement: Life Technologies, Inc.		CDNA Library Preparation: Life Technologies, Inc.
					CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		DNA Sequencing by: Incyte Genomics, Inc.
					Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:		http://image.llnl.gov
					Plate: LLAM11437 row: c column: 22		High quality sequence stop: 819.
					Location/Qualifiers		1..897
					/organism="Homo sapiens"		/mol_type="mRNA"
					/db_xref="taxon:9606"		/clone="IMAGE:5175501"
					/lab_host="DH10B"		/clone_lib="NIH_MGC_115"
					/note="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."		
			ORIGIN	Query Match	34.0%; Score 834.8; DB 2; Length 897;		Best Local Similarity
					Matches 875; Conservative 0; Mismatches 12; Indels 3; Gaps 3;		
	QY	160	GTCTCCCCAGACCTGAGGAGCTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGACCAA	219			
	Db	1	GTCTCCCCAGACCTGAGGAGCTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGACCAA	60			
	QY	220	GTTACAGCCCAACGGCTTGGCAGCTAGGCTCCAGAAATGGCTACAAAGTCCCTGATGGTTG	279			
	Db	61	GTTACAGCCCAACGGCTTGGCAGCTAGGCTCCAGAAATGGCTACAAAGTCCCTGATGGTTG	120			
	QY	280	CCGCAATGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGGCAT	339			

FEATURES	High quality sequence stop: 687.		Db	721	CACTACCTCAGATGCTCTCTCCATTGGCATGGGTGGGCTGGATCACCTGCTCAT		780
	Location/Qualifiers						
source	1..899		QY	949	GCTTCTGACTTTGACCGCAGGTGGATGACACCATCTCTAGCTCCGCTTGCTGGCAA		1008
	/organism="Homo sapiens"						
	/mol_type="mRNA"		Db	781	GCTTCTGACTTTGACCGCAGGTGGATGACACCATCTCTAGCTCCGCTTGCTGGCAA		840
	/db_xref="taxon:9606"						
	/clone="IMAGE:648086"		QY	1009	TGGCTGGG		1016
	/issue_type="carcinoma, cell line"						
	/lab_host="DH10B (phage-resistant)"		Db	841	TGGCTGGG		848
	/clone_lib="NIH_MGC_40"						
	/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		RESULT 13	BI820695	897 bp mRNA linear EST 04-OCT-2001		603034482F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175501 5', mRNA sequence.
	ACCESSION		LOCUS	DEFINITION	EST.		BI820695.1 GI:15932245
	VERSION						
	KEYWORDS		SOURCE	ORGANISM	Homo sapiens (human)		Homo sapiens
	REFERENCE		AUTHORS	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		Unpublished (1999)
	JOURNAL						
	COMMENT		FEATURES	source	Contact: Robert Strausberg, Ph.D.		Email: cgapbs-remail.nih.gov
					Tissue Procurement: Life Technologies, Inc.		CDNA Library Preparation: Life Technologies, Inc.
					CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		DNA Sequencing by: Incyte Genomics, Inc.
					Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:		http://image.llnl.gov
					Plate: LLAM11437 row: c column: 22		High quality sequence stop: 819.
					Location/Qualifiers		1..897
					/organism="Homo sapiens"		/mol_type="mRNA"
					/db_xref="taxon:9606"		/clone="IMAGE:5175501"
					/lab_host="DH10B"		/clone_lib="NIH_MGC_115"
					/note="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."		
			ORIGIN	Query Match	34.0%; Score 834.8; DB 2; Length 897;		Best Local Similarity
					Matches 875; Conservative 0; Mismatches 12; Indels 3; Gaps 3;		
	QY	160	GTCTCCCCAGACCTGAGGAGCTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGACCAA	219			
	Db	1	GTCTCCCCAGACCTGAGGAGCTCGCTGCTGCCCTCTTGGCGCGGGAAGCAGACCAA	60			
	QY	220	GTTACAGCCCAACGGCTTGGCAGCTAGGCTCCAGAAATGGCTACAAAGTCCCTGATGGTTG	279			
	Db	61	GTTACAGCCCAACGGCTTGGCAGCTAGGCTCCAGAAATGGCTACAAAGTCCCTGATGGTTG	120			
	QY	280	CCGCAATGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGGCAT	339			

Db	121	CGCAATGCGCTGAAATCAAGTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCAT	180
Qy	340	CQTCTAGAAACGGTGGCCACAGCGGGGTGTGACCTCGTGGCTTCATGCTCACTCT	399
Db	181	CQTCTAGAAACGGTGGCCACAGCGGGGTGTGACCTCGTGGCTTCATGCTCACTCT	240
Qy	400	CCCGATCTCTGCTGCAAGGTGAGGACTCCAAAGGCGGAAAAATGCTGCTACTCAGTT	459
Db	241	CCCGATCTCTGCTGCAAGGTGAGGACTCCAAAGGCGGAAAAATGCTGCTACTCAGTT	300
Qy	460	TCTCTCTCTCTGGGTGTGTGGGATCTTTGGGCTCAGCTTCGCGCTTCATCATCGGACT	519
Db	301	TCTCTCTCTCTGGGTGTGTGGGATCTTTGGGCTCAGCTTCGCGCTTCATCATCGGACT	360
Qy	520	GGACGGGAGCAGAGGGCCACACGCTTCTCTCTTTGGGATCCCTTTTCCATCTGCTTT	579
Db	361	GGACGGGAGCAGAGGGCCACACGCTTCTCTCTTTGGGATCCCTTTTCCATCTGCTTT	420
Qy	580	CTCTCTCTCTGGCTCATGCTGTGCTGACCAAGCTCGTCCGGGGAGGAAGCCCT	639
Db	421	CTCTCTCTCTGGCTCATGCTGTGCTGACCAAGCTCGTCCGGGGAGGAAGCCCT	480
Qy	640	TTCCCTCTGTGGTATCTGGGTCTGGCGTGGGCTTACGCTAGTCCAGGATGTTATCGC	699
Db	481	TTCCCTCTGTGGTATCTGGGTCTGGCGTGGGCTTACGCTAGTCCAGGATGTTATCGC	540
Qy	700	TATTGAATATATGCTGACCATGAATAGGACCAACGTCATGCTTTCTTGAGCTTTC	759
Db	541	TATTGAATATATGCTGACCATGAATAGGACCAACGTCATGCTTTCTTGAGCTTTC	600
Qy	760	CGCTCTCTGTCGAATGAAGACTTTGTCTCTCTGCTCACTACGTCCTCTTTGATGCG	819
Db	601	CGCTCTCTGTCGAATGAAGACTTTGTCTCTCTGCTCACTACGTCCTCTTTGATGCG	660
Qy	820	GCTGACCTCTCTCATGCTCTCTCTTACCTTCTGTGGTCTCTTACGCGGCTGAAGACAC	879
Db	661	GCTGACCTCTCTCATGCTCTCTCTTACCTTCTGTGGTCTCTTACGCGGCTGAAGACAC	720
Qy	880	TGGGGCCACATCTACCTCAGCATGCTCTCTCCATTGCGCATCTGGTGGCTGATCAC	939
Db	721	TGGGGCCACATCTACCTCAGCATGCTCTCTCCATTGCGCATCTGGTGGCTGATCAC	780
Qy	940	CCTGC-TCATGCTCTCTGATTTGACCGCAGGTGGGATGACACCATCTCTCAGCTCGGCT	998
Db	781	CCTGCCTCATGCTCTCTGATTTGACCGCAGGTGGGATGACACCATCTCTCAGCTCGGCT	840
Qy	999	TGGCTGCCATCGCTGGGTGTTCTGCTTGGCTATGTTAGTCCCGAGTTT	1048
Db	841	TGGCTGCC-ATGGCTGGGTG-TCCGTGTTGGTTATGTTAGTCCCGAGTTT	888
RESULT 14			
CF780868			
LOCUS			
DEFINITION			
AGENCOURT_15939245 NIH_MGC_219 Homo sapiens cDNA clone			
IMAGE:30523569 5', mRNA sequence.			
ACCESSION			
CF780868			
VERSION			
CF780868.1 GI:37740645			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1 (bases 1 to 916)			
REFERENCE			
AUTHORS			
NIH-MGC http://mgi.nci.nih.gov/.			
TITLE			
National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL			
Unpublished (1999)			
COMMENT			
Contact: Daniela S. Gerhard, Ph.D.			
Office of Cancer Genomics			
National Cancer Institute / NIH			
Bldg. 31 Rm10A07 Bethesda, MD 20892			

Email: c9apbs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM601 row: 1 column: 10
High quality sequence stop: 651.
FEATURES
source
1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30523569"
/tissue_type="Pooled Chondrosarcoma Tumor cells"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_219"
/note="Vector: pYX-Asc; Site 1: EcoRI; Site 2: NotI; Library is oligo-dT primed and directionally cloned. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 0.5-1kb. Adaptors 5' (AATTCGGCAGGAGG)3' and 5' (CTCTGCTCCG)3'. 3' Linker sequence - GGGCGCTGTGAGGCG T18. Sequencing primers 3'end: T3 promoter primer 5'd (AATTAACCTCACTAAAGGA)3'. 5' End: T7 promoter primer 5'd (TAATACGACTACTATAGG)3'. Library was constructed in the laboratory of M. Bento Soares. Average insert size 2-3kb. Note: this is a NIH_MGC Library."

ORIGIN

Query Match	33.8%;	Score	829.8;	DB	5;	Length	916;
Best Local Similarity	98.8%;	Pred. No.	2.1e-165;				
Matches	856;	Conservative	0;	Mismatches	8;	Indels	2;
Gaps	2;						
Qy	610	GACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGATTTCTGGGTCTGGCCGT	669				
Db	6	GGCCAAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGATTTCTGGGTCTGGCCGT	65				
Qy	670	GGGCTTCAGCCCTAGTCCAGGATGTTATCGCTATTTGAATATATTTGTCCTGACCATGAATAG	729				
Db	66	GGGCTTCAGCCCTAGTCCAGGATGTTATCGCTATTTGAATATATTTGTCCTGACCATGAATAG	125				
Qy	730	GACCAACGTCATGCTCTTTTCTGAGCTTTCGGCTCTCGTCGCAATGAAGACTTTGTCCT	789				
Db	126	GACCAACGTCATGCTCTTTTCTGAGCTTTCGGCTCTCGTCGCAATGAAGACTTTGTCCT	185				
Qy	790	CCTGCTCACTACGCTCTCTTCTTGATGGCGCTGACCTTCTCATGCTCTCTCTTCACTT	849				
Db	186	CCTGCTCACTACGCTCTCTTCTTGATGGCGCTGACCTTCTCATGCTCTCTCTTCACTT	245				
Qy	850	CTGTGTTCTCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACTCTCACGATGCTCT	909				
Db	246	CTGTGTTCTCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACTCTCACGATGCTCT	305				
Qy	910	CTCCATTTGCCATCTGGTGGCTTGGATCACCCCTGCTCATGCTTCTCTGACTTTGACCGCAG	969				
Db	306	CTCCATTTGCCATCTGGTGGCTTGGATCACCCCTGCTCATGCTTCTCTGACTTTGACCGCAG	365				
Qy	970	GTGGGATGACACATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGTCTCTGTGGC	1029				
Db	366	GTGGGATGACACATCTCAGCTCCGCTTGGCTGGCAATGGCTGGGTGTCTCTGTGGC	425				
Qy	1030	TTATGTTAGTCCCGAGTTTGGCTGCTCAAAAGCAACGAAACCCCATGGATTATCTGT	1089				
Db	426	TTATGTTAGTCCCGAGTTTGGCTGCTCAAAAGCAACGAAACCCCATGGATTATCTGT	485				
Qy	1090	TGAGGATGCTTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGGAGAACAGAGC	1149				


```
Db 486 TGAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGC 545
Qy 1150 CTACTCTCAAGAGAAATCACTCAAGTTTGAAGAGACAGGGGACACGCTCTATGCCCC 1209
Db 546 CTACTCTCAAGAGAAATCACTCAAGTTTGAAGAGACAGGGGACACGCTCTATGCCCC 605
Qy 1210 CTATTCACACATTTTCAGTCGAGAACAGACGCTCCCAAAAGAAATTTCCATCCCAAG 1269
Db 606 CTATTCACACATTTTCAGTCGAGAACAGACGCTCCCAAAAGAAATTTCCATCCCAAG 665
Qy 1270 GGGCCAGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAGAAAGAGGCGAGCTAACT 1329
Db 666 GGGCCAGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAGAAAGAGGCGAGCTAACT 725
Qy 1330 CTGCTCTGAAGAGTGGGACAAATGACGCGGGGCGGAGACTCTAGCGGGAGCTCAAGAGGA 1389
Db 726 CTGCTCTGAAGAGTGGGACAAATGACGCGGGGCGGAGACTCTAGCGGGAGCTCAAGAGGA 785
Qy 1390 TGTGGCGGAAATCTTCAGTCTTCTGAGAAACTGTACAGACACTACGGGAACAG-TTTG 1448
Db 786 TGTGGCGGAAATCTTCAGTCTTCTGAG-TAACTGTACAGACACTACGGGAACCGTTTGG 844
Qy 1449 CTCTCCCTCCAGCTCAACACAAAT 1474
Db 845 CTCTCCCTCCAGCTCAACACAAAT 870
```

RESULT 15

```
AL546138
LOCUS AL546138 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1025YF10 5-PRIME, mRNA sequence.
ACCESSION AL546138
VERSION .AL546138.3 GI:45746606
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31267972.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5601.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1025DC050Pl&c=6601.f.
```

FEATURES

```
source
1..839
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1025YF10"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 33.4%; Score 820; DB 1; Length 839;

```
Best Local Similarity 99.2%; Pred. No. 2.5e-163;
Matches 823; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
Qy 160 GTCTCCCCCAGACATGAGGAGCTCGCTGCTGCTTTCGGCGCGGGAAGACAGACCAA 219
Db 1 GTCTCCCCCAGACATGAGGAGCTCGCTGCTGCTTTCGGCGCGGGAAGACAGACCAA 60
Qy 220 GTTTCAGCGGCAACGCGCTTCGGCACTAGGGTTCAGAAATGGCTTACAAAGTCCCTCATGGTTG 279
Db 61 GTTTCAGCGGCAACGCGCTTCGGCACTAGGGTTCAGAAATGGCTTACAAAGTCCCTCATGGTTG 120
Qy 280 CCGCAATGCGCTGAAATCCAAAGTACTACAGACTTTTGTGATAAGGCTTGAAGCTTTGGGCGCAT 339
Db 121 CCGCAATGCGCTGAAATCCAAAGTACTACAGACTTTTGTGATAAGGCTTGAAGCTTTGGGCGCAT 180
Qy 340 CGTCTCTAGAAAAGTGGTGGCCACAGCGGGGTGTGACCTCGGTGGCTTCATGCTCACTCT 399
Db 181 CGTCTCTAGAAAAGTGGTGGCCACAGCTGGGGTGTGACCTCAGTGGCTTCATGCTCACTCT 240
Qy 400 CCGCATCTCTGCTGCAAGGTGCAGGACTCCAAACAGCGGCAAAATGCTGCCTACTCAGTT 459
Db 241 CCGCATCTCTGCTGCAAGGTGCAGGACTCCAAACAGCGGCAAAATGCTGCCTACTCAGTT 300
Qy 460 TCTCTTCTCTCTGGGTGTGTGGGCACTTTTGGCCCTCACCTTCGCCCTTCATCATCGGACT 519
Db 301 TCTCTTCTCTCTGGGTGTGTGGGCACTTTTGGCCCTCACCTTCGCCCTTCATCATCGGACT 360
Qy 520 GGACGGGAGACACAGGGCCCAACGCTTCTCTCTTTGGGATCCTCTTTTCAATCTGCTT 579
Db 361 GGACGGGAGACACAGGGCCCAACGCTTCTCTCTTTGGGATCCTCTTTTCAATCTGCTT 420
Qy 580 CTCCTCGCTCTGGGCTCATGCTGTGCTGACCTGACCAAGCTCGTCCGGGGAGGAGCCCT 639
Db 421 CTCCTCGCTCTGGGCTCATGCTGTGCTGACCAAGCTCGTCCGGGGAGGAGCCCT 480
Qy 640 TTCCCTCTGTTGGTGAATTTCTGGGCTTGGCGCTTGGGCTTTCAGCTAGTCCAGGATGTTATCGC 699
Db 481 TTCCCTCTGTTGGTGAATTTCTGGGCTTGGCGCTTGGGCTTTCAGCTAGTCCAGGATGTTATCGC 540
Qy 700 TATTGAATATATGCTCTGACCATGAATAGGACCAACGTCATGCTTTCTGAGCTTTC 759
Db 541 TATTGAATATATGCTCTGACCATGAATAGGACCAACGTCATGCTTTCTGAGCTTTC 600
Qy 760 CGCTCTCTGCTGCAATGAAGACTTTGTCCTCTGCTGCTCAGCTAGCTCTCTTTGATGGC 819
Db 601 CGCTCTCTGCTGCAATGAAGACTTTGTCCTCTGCTGCTCAGCTAGCTCTCTTTGATGGC 660
Qy 820 GCTGACCTTCTCTCATGCTCCTCTTCACTTCTGCTGCTTCTTTCACGGGCTTGAAGAGACA 879
Db 661 GCTGACCTTCTCTCATGCTCCTCTTCACTTCTGCTGCTTCTTTCACGGGCTTGAAGAGACA 720
Qy 880 TGGGGCCCAATCTACCTCAAGTCTCTCTCCATGTCATGCTGGGTGGCTGGATCAC 939
Db 721 TGGGGCCCAATCTACCTCAAGTCTCTCTCCATGTCATGCTGGGTGGCTGGATCAC 780
Qy 940 CTTGCTCATGCTTCTGACTTTGACGGGAGTGGGATGACACCATCTCA 989
Db 781 CTTGCTCATGCTTCTGACTTTGACGGGAGTGGGATGACACCATCTCA 830
```

Search completed: June 10, 2006, 18:50:04

Job time : 11520 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:02:15 ; Search time 13877 Seconds
(without alignments)
11317.638 Million cell updates/sec

Title: US-10-600-816-2
Perfect score: 2456
Sequence: 1 ataacagcatgaagtcgcgt.....ggcagcaaaaaaaaaa 2456

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: g5_vi.*
11: gb_ov.*
12: gb_htg.*
13: gb_in.*
14: gb_om.*
15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2456	100.0	2456	2	CQ981495	CQ981495 Sequence
2	2456	100.0	2456	2	DD210040	DD210040 Methods o
3	2456	100.0	2456	2	AX549168	AX549168 Sequence
4	2443	99.5	6730	2	CQ894732	CQ894732 Sequence
5	2439.8	99.3	2446	2	BD156680	BD156680 Primer fo
6	2439.8	99.3	2446	2	AX877483	AX877483 Sequence
7	2439.8	99.3	2446	5	AK001761	AK001761 Homo sapi
8	2418.4	98.5	3057	5	AK122672	AK122672 Homo sapi
9	2302	93.7	2302	2	AX930411	AX930411 Sequence
10	2302	93.7	2302	5	AF095448	AF095448 Homo sapi
11	2288.8	93.2	3371	2	AX188348	AX188348 Sequence
12	2274.8	92.6	2290	2	CQ723177	CQ723177 Sequence
13	2274.8	92.6	2297	5	AF506289	AF506289 Homo sapi
14	2266.4	92.3	2296	5	BC003665	BC003665 Homo sapi
15	1795.4	73.1	1826	5	AK172760	AK172760 Homo sapi
16	1701.2	69.3	1718	2	BD233473	BD233473 Human pro
17	1601.8	65.2	1619	2	AX078375	AX078375 Sequence
18	1215.4	49.5	161577	5	AC007688	AC007688 Homo sapi

19	1126	45.8	1212	2	BD209699	BD209699 Compositi
20	1126	45.8	1212	2	AR341505	AR341505 Sequence
21	1082.6	44.1	1114	2	BD139393	BD139393 Extended
C 22	932.6	38.0	1000	2	CQ056634	CQ056634 Sequence
C 23	932.6	38.0	1000	2	CQ075911	CQ075911 Sequence
C 24	932.6	38.0	1000	2	CQ106888	CQ106888 Sequence
C 25	932.6	38.0	1000	2	CQ145543	CQ145543 Sequence
C 26	932.6	38.0	1000	2	CQ180988	CQ180988 Sequence
C 27	932.6	38.0	1000	2	CQ205353	CQ205353 Sequence
C 28	932.6	38.0	1000	2	CQ228751	CQ228751 Sequence
C 29	932.6	38.0	1000	2	CQ266894	CQ266894 Sequence
C 30	932.6	38.0	1000	2	CQ303866	CQ303866 Sequence
C 31	932.6	38.0	1000	2	CQ341182	CQ341182 Sequence
C 32	815.2	33.2	2069	14	BC105147	BC105147 Bos tauru
C 33	787.6	32.1	948	2	CQ051666	CQ051666 Sequence
C 34	787.6	32.1	948	2	CQ066723	CQ066723 Sequence
C 35	787.6	32.1	948	2	CQ093765	CQ093765 Sequence
C 36	787.6	32.1	948	2	CQ132542	CQ132542 Sequence
C 37	787.6	32.1	948	2	CQ171114	CQ171114 Sequence
C 38	787.6	32.1	948	2	CQ200250	CQ200250 Sequence
C 39	787.6	32.1	948	2	CQ215753	CQ215753 Sequence
C 40	787.6	32.1	948	2	CQ254332	CQ254332 Sequence
C 41	787.6	32.1	948	2	CQ291365	CQ291365 Sequence
C 42	787.6	32.1	948	2	CQ328468	CQ328468 Sequence
C 43	708.4	28.8	2102	6	BC039217	BC039217 Mus muscu
C 44	708.4	28.8	2174	6	BC036173	BC036173 Mus muscu
C 45	668.4	27.2	1936	6	BC036174	BC036174 Mus muscu

ALIGNMENTS

RESULT 1
CQ981495
LOCUS CQ981495 2456 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 350 from Patent EP1498424.
ACCESSION CQ981495
VERSION CQ981495.1 GI:58190785
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Bruemendorf, T., Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE Human nucleic acid sequences from lung tumours
JOURNAL Patent: EP 1498424-A 350 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE); Rosenthal, Andre (DE)

FEATURES
source Location/Qualifiers
1..2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2456; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATACAGCATGAGTCCGCTGGAACCTGGAATAGCGGTGCTCTCCCTCCGAGGGCGGCTTTATA 120
DB 1 ATACAGCATGAGTCCGCTGGAACCTGGAATAGCGGTGCTCTCCCTCCGAGGGCGGCTTTATA 120
QY 61 TCCTTCTCCCTCTGCTCACCCTCGCTTCCTCCCTCCGAGGGCGGCTTTATA 120
DB 61 TCCTTCTCCCTCTGCTCACCCTCGCTTCCTCCCTCCGAGGGCGGCTTTATA 120
QY 121 ACAACTGCTCAGAGTCGAGGGCGGATAGCTCCAAAGGTCTCCCGAGCTAGGAG 180
DB 121 ACAACTGCTCAGAGTCGAGGGCGGATAGCTCCAAAGGTCTCCCGAGCTAGGAG 180

Qy	181	CTCGCTCTGTCCTCTTTGGCGCGGGAAGCAGCAACAAGTTACAGGCCAACGCCTTGGC	240
Db	181	CTCGCTGCTGTCCTCTTTGGCGCGGGAAGCAGCAACAAGTTACAGGCCAACGCCTTGGC	240
Qy	241	ACTAGGGTCAGANTGGCTACAAACAGTCCCTGATGGTTGGCGAATGGCTGAATCCAA	300
Db	241	ACTAGGGTCAGAAATGGCTACAAACAGTCCCTGATGGTTGGCGAATGGCTGAATCCAA	300
Qy	301	GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360
Db	301	GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATCGTCTAGAAACGGTGGCCAC	360
Qy	361	AGCGGGGTTGTGACTCTCGTGGGCTTATGCTCTCACTCTCCGATCTCGTCTGCAAGGT	420
Db	361	AGCGGGGTTGTGACTCTCGTGGGCTTATGCTCTCACTCTCCGATCTCGTCTGCAAGGT	420
Qy	421	GCAGGACTCCAAACAGGCGAAATATGCTGCTACTCAGTTTCTCTTCCTCTCGTGGTGTGT	480
Db	421	GCAGGACTCCAAACAGGCGAAATATGCTGCTACTCAGTTTCTCTTCCTCTCGTGGTGTGT	480
Qy	481	GGGCATCTTTGGGCTCTACCTTTGCGCTTCACTCGGACTGGACGGGAGCAACGGGCCAC	540
Db	481	GGGCATCTTTGGGCTCTACCTTTGCGCTTCACTCGGACTGGACGGGAGCAACGGGCCAC	540
Qy	541	ACGCTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCCTGCTGCTGCTCATGC	600
Db	541	ACGCTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCCTGCTGCTGCTCATGC	600
Qy	601	TGTCAGTCTGACCAAGCTCGTCCGGGGGAGGAAGCCCTTTTCCCTGTTGGTGAATCTGGG	660
Db	601	TGTCAGTCTGACCAAGCTCGTCCGGGGGAGGAAGCCCTTTTCCCTGTTGGTGAATCTGGG	660
Qy	661	TCTGGCGTGGGCTTCAAGCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGAC	720
Db	661	TCTGGCGTGGGCTTCAAGCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGAC	720
Qy	721	CATGAATAGGACCAACGTCGAATGCTTTTCTGAGCTTTTCCGCTCTCGTCGCAATGAAGA	780
Db	721	CATGAATAGGACCAACGTCGAATGCTTTTCTGAGCTTTTCCGCTCTCGTCGCAATGAAGA	780
Qy	781	CTTTGTCTCTGCTGCTCACTTCTTGATGGCGCTGACCTTCCTCATGTCCTC	840
Db	781	CTTTGTCTCTGCTGCTCACTTCTTGATGGCGCTGACCTTCCTCATGTCCTC	840
Qy	841	CTTCACTCTTGTGGTTCCTTCAAGGCTGGAGAGACATGGGGCCACATCTACCTCAC	900
Db	841	CTTCACTCTTGTGGTTCCTTCAAGGCTGGAGAGACATGGGGCCACATCTACCTCAC	900
Qy	901	GATGCTCTCTCCATTCGCATCTGGGTGGCTGGATCACCCCTGCTCATGCTTCTGACTT	960
Db	901	GATGCTCTCTCCATTCGCATCTGGGTGGCTGGATCACCCCTGCTCATGCTTCTGACTT	960
Qy	961	TGACCGCAGGTGGATGACACCATCTCAGCTCGCCTTGGCTGCCAATGGCTGGGTGT	1020
Db	961	TGACCGCAGGTGGATGACACCATCTCAGCTCGCCTTGGCTGCCAATGGCTGGGTGT	1020
Qy	1021	CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAAAAGCAAACGAAACCCCATGGA	1080
Db	1021	CCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAAAAGCAAACGAAACCCCATGGA	1080
Qy	1081	TTATCTCTGTTGAGGATGTTCTGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGA	1140
Db	1081	TTATCTCTGTTGAGGATGTTCTGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGA	1140
Qy	1141	GAAACAGGCTTACTCTCAAGAGGAAATCCTCAAGGTTTTTGAAGAGACAGGGACACGCT	1200
Db	1141	GAAACAGGCTTACTCTCAAGAGGAAATCCTCAAGGTTTTTGAAGAGACAGGGACACGCT	1200
Qy	1201	CTATGCCCCCTATTCCACACATTTTACGCTGCAGAAACGACCTCCCCAAAGGAATTTCTC	1260
Db	1201	CTATGCCCCCTATTCCACACATTTTACGCTGCAGAAACGACCTCCCCAAAGGAATTTCTC	1260

Qy	1261	CATCCACGCGGCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAGAAAGAGGG	1320
Db	1261	CATCCACGCGGCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAGAAAGAGGG	1320
Qy	1321	CAGCTAACTCTGTCTGAAGAGTGGACAAATCAGCGGGCGGCAGATCTACGGGAGC	1380
Db	1321	CAGCTAACTCTGTCTGAAGAGTGGACAAATCAGCGGGCGGCAGATCTACGGGAGC	1380
Qy	1381	TCAAGGAGTGTGGGCGAAATCTTGAGTCTTCTTGAGAAAACTGTACAAGACACTACGGGA	1440
Db	1381	TCAAGGAGTGTGGGCGAAATCTTGAGTCTTCTTGAGAAAACTGTACAAGACACTACGGGA	1440
Qy	1441	ACAGTTGGCTCCCTCCAGCCTCAACCAACAATCTTCCAGTGTGGGGCTGATGTGGGCT	1500
Db	1441	ACAGTTGGCTCCCTCCAGCCTCAACCAACAATCTTCCAGTGTGGGGCTGATGTGGGCT	1500
Qy	1501	AGTAAGACTCCAGTCTTTAGAGCGGTGTAGTATTTTTTTTTTTTGTCTCATCCTTTGG	1560
Db	1501	AGTAAGACTCCAGTCTTTAGAGCGGTGTAGTATTTTTTTTTTTGTCTCATCCTTTGG	1560
Qy	1561	ATACTTCTTTTAAGTGGGAGTCTCAGCAAACTCAAGTTTAGACCTTACTCTTTTGTGTTT	1620
Db	1561	ATACTTCTTTTAAGTGGGAGTCTCAGCAAACTCAAGTTTAGACCTTACTCTTTTGTGTTT	1620
Qy	1621	GTTTTTTGAACAGGATCTTGCTCTGTCAACCAGGCTTGAGTCAGTGGTGCATCACAG	1680
Db	1621	GTTTTTTGAACAGGATCTTGCTCTGTCAACCAGGCTTGAGTCAGTGGTGCATCACAG	1680
Qy	1681	CCAGTGCAGCCTCGACCACTGTGCTCAAGCAAACTCTCCCATCTCCATCTCCCAAGTG	1740
Db	1681	CCAGTGCAGCCTCGACCACTGTGCTCAAGCAAACTCTCCCATCTCCATCTCCCAAGTG	1740
Qy	1741	CTGGATGACAGCGGTGAGCCACAGCTCCAGCCTAGGCCCTTAACTCTGCTGTATTTT	1800
Db	1741	CTGGATGACAGCGGTGAGCCACAGCTCCAGCCTAGGCCCTTAACTCTGCTGTATTTT	1800
Qy	1801	CCATGGACTAAAGGCTCTGGTCACTCAGCTCAGCTGGCTCACAAGCTCTAGGGGCGTG	1860
Db	1801	CCATGGACTAAAGGCTCTGGTCACTCAGCTCAGCTGGCTCACAAGCTCTAGGGGCGTG	1860
Qy	1861	CTCCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACTGCATATCT	1920
Db	1861	CTCCTCTAACTCAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACTGCATATCT	1920
Qy	1921	GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCTGAATCTACACTGGGAAGCCAAC	1980
Db	1921	GAGCAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCTGAATCTACACTGGGAAGCCAAC	1980
Qy	1981	TTGCTGCGACCCCGCTCCCAACCTCTTCTGCTGGGTAGGAGAGGCTAAAGATCAACC	2040
Db	1981	TTGCTGCGACCCCGCTCCCAACCTCTTCTGCTGGGTAGGAGAGGCTAAAGATCAACC	2040
Qy	2041	TAAATTTACTCATCTCTCTAGTGTGCTCACAATTGGGCTCTCAGCAGCTCCCGACACCA	2100
Db	2041	TAAATTTACTCATCTCTCTAGTGTGCTCACAATTGGGCTCTCAGCAGCTCCCGACACCA	2100
Qy	2101	ATTCACAGGTCACCCCTCTCTTCTTGCACTGTCCCAAACTTGCTGTCAAATCCGAGATC	2160
Db	2101	ATTCACAGGTCACCCCTCTCTTCTTGCACTGTCCCAAACTTGCTGTCAAATCCGAGATC	2160
Qy	2161	TAAATCTCCCTCAGCTCTCGAGGATCTTTTCAGACCTCACTAGCACAAAGCCCGGTG	2220
Db	2161	TAAATCTCCCTCAGCTCTCGAGGATCTTTTCAGACCTCACTAGCACAAAGCCCGGTG	2220
Qy	2221	CTCCTTGTCAAGGAAATTTGTAGATCATTTCTCAATCAAAATTCCTGGGGCTGATCTTCT	2280
Db	2221	CTCCTTGTCAAGGAAATTTGTAGATCATTTCTCAATCAAAATTCCTGGGGCTGATCTTCT	2280
Qy	2281	CTCATCTTGACCCCAACCTCTGTAAATAGATTTACGGCATTTACGGCTGCATCTGTAA	2340
Db	2281	CTCATCTTGACCCCAACCTCTGTAAATAGATTTACGGCATTTACGGCTGCATCTGTGTAA	2340
Qy	2341	GTGGGCATGTCTCCTAATGGAGAGTGTTTCATTGTATATAAGTTATTCACTCTGAGTAT	2400

```
|||||
Db 2341 GTGGCGATGGTCTCTTAATGAGAGAGTTCATTGTATATAAGTATTACCTCGACTAT 2400
Qy 2401 GCAATAAAGATGGTGGCCACTCTTTTCATGGTGGTGGGAGCAAAAAA 2456
Db 2401 GCAATAAAGATGGTGGCCACTCTTTTCATGGTGGTGGGAGCAAAAAA 2456

RESULT 2
DD210040 2456 bp DNA linear PAT 19-JAN-2006
LOCUS Methods of Diagnosis of Cancer, Compositions and Methods of
DEFINITION Screening for Modulators of Cancer.
ACCESSION DD210040
VERSION DD210040.1 GI-85654022
KEYWORDS JP 2005518782-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2456)
AUTHORS Zlotnik,A., Mack,D.H., Agiz,N., Gish,K.C., Hebeji,P.A., Wilson,K.E.
and Afar,D.
TITLE Methods of Diagnosis of Cancer, Compositions and Methods of
Screening for Modulators of Cancer
JOURNAL Patent: JP 2005518782-A 34 30-JUN-2005;
PROTEIN DESIGN LABS INC
COMMENT OS Homo sapiens
PN JP 2005518782-A/34
PD 30-JUN-2005
PF 17-SEP-2002 JP 2003529912
PR 12-APR-2002 US 60/372246, 08-FEB-2002 US 60/355257, PR
08-FEB-2002 US 60/355145, 13-NOV-2001 US 60/350666, PR
20-SEP-2001 US 60/323887, 17-SEP-2001 US 60/323469 PI
Zlotnik,david h mack,natasha agiz,kurt c gish,peter a pi
PI keith e wilson,daniel afar
CC
FH Key Location/Qualifiers.
source 1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 2456; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Qy 1 ATACAGCATGAGTGGCGTGGAACTGGAATAGCGGTGCTCTCCCTCGACCTCCCC 60
Db 1 ATACAGCATGAGTGGCGTGGAACTGGAATAGCGGTGCTCTCTCCCTCGACCTCCCC 60
Qy 61 TCCTTGTCCTCTGCTCACCCCTCGCTGTTTCCCTCCCTCGGGAGGGCGGCTTTATA 120
Db 61 TCCTTGTCCTCTGCTCACCCCTCGCTGTTTCCCTCCCTCGGGAGGGCGGCTTTATA 120
Qy 121 ACAACTGCTCAGAGTGGAGGGGGATAGCTGTCCAAAGTCTCCCCAGCACTGAGGAG 180
Db 121 ACAACTGCTCAGAGTGGAGGGGGATAGCTGTCCAAAGTCTCCCCAGCACTGAGGAG 180
Qy 181 CTGCGCTGCTGCTCTTGGCGGGGAGGAGCAGCACCAGTTTACGGCCCAACGGCTTGGC 240
Db 181 CTGCGCTGCTGCTCTTGGCGGGGAGGAGCAGCACCAGTTTACGGCCCAACGGCTTGGC 240
Qy 241 ACTAGGTCCAGATGCTTACCAAGTCCCTGATGTTGGCGCAATGGCTGAAATCCAA 300
Db 241 ACTAGGTCCAGATGCTTACCAAGTCCCTGATGTTGGCGCAATGGCTGAAATCCAA 300
Qy 301 GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATGCTCTAGAAAACGGTGGCCAC 360
Db 301 GTACTACAGACTTTGTGATAGGCTGAAGCTTGGGGCATGCTCTAGAAAACGGTGGCCAC 360
```

```
Qy 361 AGCGGGGTGTGACCTCGTGGCCTTCACTGCTCACTCTCCCGATCCTCTCTGCAAGGT 420
Db 361 AGCGGGGTGTGACCTCGTGGCCTTCACTGCTCACTCTCCCGATCCTCTCTGCAAGGT 420
Qy 421 GCAGGACTCCAAACAGGGCAAAAATGTCCTACTCAGATTTCTCTCTCTCTCTGCTGTT 480
Db 421 GCAGGACTCCAAACAGGGCAAAAATGTCCTACTCAGTTTCTCTCTCTCTCTGCTGTT 480
Qy 481 GGGCATCTTTGGCTCACCTTCCCTTCACTCGGACTCGGACGGGAGCAACAGGGCCAC 540
Db 481 GGGCATCTTTGGCTCACCTTCCCTTCACTCGGACTCGGACGGGAGCAACAGGGCCAC 540
Qy 541 AGCTTCTTCTCTCTTTGGGATCTCTTTTCCATCTCTCTCTCTCTCTCTCTCTCTGTC 600
Db 541 AGCTTCTTCTCTCTTTGGGATCTCTTTTCCATCTCTCTCTCTCTCTCTCTCTCTGTC 600
Qy 601 TGTGAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGAATTCGGG 660
Db 601 TGTGAGTCTGACCAAGCTCGTCCGGGGAGGAGCCCTTTCCCTGTTGGTGAATTCGGG 660
Qy 661 TCTGGCGGTGGCTTCAAGCTAGTCAGGATGTTATCGCTATTGAATATATTGTCCTGAC 720
Db 661 TCTGGCGGTGGCTTCAAGCTAGTCAGGATGTTATCGCTATTGAATATATTGTCCTGAC 720
Qy 721 CATGAATAGGACCAAGCTCAATGCTCTTTTCTGAGCTTTCCGCTCTCGCTGCAATGAAGA 780
Db 721 CATGAATAGGACCAAGCTCAATGCTCTTTTCTGAGCTTTCCGCTCTCGCTGCAATGAAGA 780
Qy 781 CTTTGTCCTCTGCTCACCTACCTCTCTTTTGTGATGGGCTGACCTTCTCTCATGTCTCTC 840
Db 781 CTTTGTCCTCTGCTCACCTACCTCTCTTTTGTGATGGGCTGACCTTCTCTCATGTCTCTC 840
Qy 841 CTTCACTCTCTGTTGGTCTCTTCAAGGCTGGAAAGACATAGGGGCCACATCTACTCTCAC 900
Db 841 CTTCACTCTCTGTTGGTCTCTTCAAGGCTGGAAAGACATAGGGGCCACATCTACTCTCAC 900
Qy 901 GATGCTCTCTCCATGTCATCTGGGTGGCTGGATCACCCTCTCATGCTCTCTGACTT 960
Db 901 GATGCTCTCTCCATGTCATCTGGGTGGCTGGATCACCCTCTCATGCTCTCTGACTT 960
Qy 961 TGACCGCAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCGCAATGGCTGGTGT 1020
Db 961 TGACCGCAGTGGGATGACACCATCTCAGCTCCGCTTGGCTGCGCAATGGCTGGTGT 1020
Qy 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
Db 1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA 1080
Qy 1081 TTATCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGTTGTGGA 1140
Db 1081 TTATCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGTTGTGGA 1140
Qy 1141 GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTTTGAAGAGACAGGGACAGCT 1200
Db 1141 GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTTTGAAGAGACAGGGACAGCT 1200
Qy 1201 CTATGCCCCCTATTTCACACATTTTTCAGCTGAGAAACAGCTCTCCCAAGAAATTTCTC 1260
Db 1201 CTATGCCCCCTATTTCACACATTTTTCAGCTGAGAAACAGCTCTCCCAAGAAATTTCTC 1260
Qy 1261 CATCCCAAGGGCCCGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG 1320
Db 1261 CATCCCAAGGGCCCGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG 1320
Qy 1321 CAGCTAACTCTGCTGAAAGAGTGGGACAAAATGACGCGGGCGGCGAGATCTAGCGGAGC 1380
Db 1321 CAGCTAACTCTGCTGAAAGAGTGGGACAAAATGACGCGGGCGGCGAGATCTAGCGGAGC 1380
Qy 1381 TCAAGGGATGTGGGCGAAATCTTGAGTCTTTCTGAGAAAACCTGTAAGAACACTACGGGA 1440
Db 1381 TCAAGGGATGTGGGCGAAATCTTGAGTCTTTCTGAGAAAACCTGTAAGAACACTACGGGA 1440
```

QY 1441 ACAGTTTGCTCCCTCCAGCCTCAACCAAAATCTTCAATGCTGGGCTGATGTGGGT 1500
DB 1441 ACAGTTTGCTCCCTCCAGCCTCAACCAAAATCTTCAATGCTGGGCTGATGTGGGT 1500
QY 1501 AGTAAGACTCCAGTCTTGAAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCCTTTGG 1560
DB 1501 AGTAAGACTCCAGTCTTGAAGCGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCCTTTGG 1560
QY 1561 ATACTTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
DB 1561 ATACTTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCTTACTCTTTTGT 1620
QY 1621 GTTTTTTGAACAGGATCTGTCTGTCAACAGGCTTGAGTCAGTGGGATCACAG 1680
DB 1621 GTTTTTTGAACAGGATCTGTCTGTCAACAGGCTTGAGTCAGTGGGATCACAG 1680
QY 1681 CCAGTGCAGCCTCGACCTGTCTGTCAAGCAATCTCCATCTCCATCTCCCAAAGTG 1740
DB 1681 CCAGTGCAGCCTCGACCTGTCTGTCAAGCAATCTCCATCTCCATCTCCCAAAGTG 1740
QY 1741 CTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTAGGCGCTTAAATCTTCTGTATTTT 1800
DB 1741 CTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTAGGCGCTTAAATCTTCTGTATTTT 1800
QY 1801 CCATGGAATTAAGGTCTGTCTGTCAAGCTCAGCTGGCTCAGACGCTTAGGGGCGTG 1860
DB 1801 CCA TGGAATAAAGGTCTGTCTGTCAAGCTCAGCTGGCTCAGACGCTTAGGGGCGTG 1860
QY 1861 CTCCTCTAACTCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
DB 1861 CTCCTCTAACTCAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
QY 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCACTGGCCTGAATCTACACTGGAAGCAAC 1980
DB 1921 GAGCAAAAATAGCAAAAGCCTCTCTCAGCCACTGGCCTGAATCTACACTGGAAGCAAC 1980
QY 1981 TTGCTGGCACCCTGCTCCCAACCTCTCTGCTGGGTAGGAGGCTAAGATCACCC 2040
DB 1981 TTGCTGGCACCCTGCTCCCAACCTCTCTGCTGGGTAGGAGGCTAAGATCACCC 2040
QY 2041 TAAATTTACTCATCTCTAGTGTCTGCCTCACAATGGGCTCAGCAGCTCCCAAGCA 2100
DB 2041 TAAATTTACTCATCTCTAGTGTCTGCCTCACAATGGGCTCAGCAGCTCCCAAGCA 2100
QY 2101 ATTACAGGTACCCCTCTCTTTGCACTGTCCCAAACTTGCTGTCAATTCGAGATC 2160
DB 2101 ATTACAGGTACCCCTCTCTTTGCACTGTCCCAAACTTGCTGTCAATTCGAGATC 2160
QY 2161 TAATCTCCCTAGCTCTGCAAGGAAATCTTTTACAGCCTCAGTACAGCAAGCCGGTTG 2220
DB 2161 TAATCTCCCTAGCTCTGCAAGGAAATCTTTTACAGCCTCAGTACAGCAAGCCGGTTG 2220
QY 2221 CTCTTTGTGAGGAAATTTGTAGATCAATCTCACTTCAAAATCTGGGCTGATCTTCT 2280
DB 2221 CTCTTTGTGAGGAAATTTGTAGATCAATCTCACTTCAAAATCTGGGCTGATCTTCT 2280
QY 2281 CTCAATCTTGCACCCCAACCCTCTGTAATAGATTACCGCAATTTACCGCTGCAATCTGTAA 2340
DB 2281 CTCAATCTTGCACCCCAACCCTCTGTAATAGATTACCGCAATTTACCGCTGCAATCTGTAA 2340
QY 2341 GTGGGATGCTCTTAATGGAGAGTGTTCATTTGATATTAATAGTTTATTCACCTGAGTAT 2400
DB 2341 GTGGGATGCTCTTAATGGAGAGTGTTCATTTGATATTAATAGTTTATTCACCTGAGTAT 2400
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456
DB 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA 2456

RESULT 3
AX549168 2456 bp DNA linear PAT 26-NOV-2002
LOCUS
DEFINITION Sequence 453 from Patent WO02061087.

ACCESSION AX549168
VERSION AX549168.1 GI:25813894
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Burner,G.C., Roush,C.L. and Brown,J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
JOURNAL Patent: WO 02061087-A 453 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2456; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAACAGCATGAAGTGCCTGGAACCTGGAATAGGCGTGTCTCTCCCTGCAACCTCCCC 60
DB 1 ATAAACAGCATGAAGTGCCTGGAACCTGGAATAGGCGTGTCTCTCCCTGCAACCTCCCC 60
QY 61 TCTTTGTCCCTCTGCTCACCCCTCGCTCCCTCCCTCCCTCCGCGAGGCGCGCTTTATA 120
DB 61 TCTTTGTCCCTCTGCTCACCCCTCGCTCCCTCCCTCCCTCCGCGAGGCGCGCTTTATA 120
QY 121 ACAACTGCTCTCAGAGTGCAGAGCGGCGGATAGTGTCTCAAGGTCTCCCCAGCACTGAGGAG 180
DB 121 ACAACTGCTCTCAGAGTGCAGAGCGGCGGATAGTGTCTCAAGGTCTCCCCAGCACTGAGGAG 180
QY 181 CTCGCTGTGCTCCTTTGCGGCGGGAAGCAGCACCAAGTTTCACGGCCAAACGCTTGGC 240
DB 181 CTCGCTGTGCTCCTTTGCGGCGGGAAGCAGCACCAAGTTTCACGGCCAAACGCTTGGC 240
QY 241 ACTAGGCTCCAGATGGCTTACACAGTCTGATGTTCGCGCAATGGCTGAAATCCAA 300
DB 241 ACTAGGCTCCAGATGGCTTACACAGTCTGATGTTCGCGCAATGGCTGAAATCCAA 300
QY 301 GTACTACAGACTTTGTGATAAGCTTTGGGCACTGTCTAGAAAACGCTGGCCAC 360
DB 301 GTACTACAGACTTTGTGATAAGCTTTGGGCACTGTCTAGAAAACGCTGGCCAC 360
QY 361 AGCGGGGTGTGACCTCGGTGCGCTTATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420
DB 361 AGCGGGGTGTGACCTCGGTGCGCTTATGCTCACTCTCCGATCCTCGTCTGCAAGGT 420
QY 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTGGGTGTGT 480
DB 421 GCAGGACTCCAAACAGCGGAAATGCTGCTACTCAGTTTCTCTCTCTGGGTGTGT 480
QY 481 GGGCATCTTTGGCCTCACCTTGGCCTTATCATCGGACTGGAGCGGAGCACAGGGCCAC 540
DB 481 GGGCATCTTTGGCCTCACCTTGGCCTTATCATCGGACTGGAGCGGAGCACAGGGCCAC 540
QY 541 AGCTTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTGCTGCTATGC 600
DB 541 AGCTTTCTTCTCTTTGGGATCTCTTTTCCATCTGCTTCTCTGCTGCTGCTATGC 600
QY 601 TGTCACTGTACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCCTGTTGGTGTATTTCTGGG 660
DB 601 TGTCACTGTACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCCTGTTGGTGTATTTCTGGG 660
QY 661 TCTGGCGGTGGGCTTCAGCCTAGTCCAGGATGTTATCGCTATTTGAATATATTTCTCTGAC 720
DB 661 TCTGGCGGTGGGCTTCAGCCTAGTCCAGGATGTTATCGCTATTTGAATATATTTCTCTGAC 720

Qy	721	CATGAATAGGACCAACGTCATATGCTCTTTCTGAGCTTTTCGGCTCCTCTCGTGGCAATGAAGA	780
Db	721	CATGARAATAGGACCAACGTCATATGCTCTTTCTGAGCTTTTCGGCTCCTCTCGTGGCAATGAAGA	780
Qy	781	CTTTGTCTCTCTGCTCACCTACGTCCTCTCTTTCTGATGGCGTGACCTTTCTCTCATGTCCCTC	840
Db	781	CTTTGTCTCTCTGCTCACCTACGTCCTCTCTTTCTGATGGCGTGACCTTTCTCTCATGTCCCTC	840
Qy	841	CTTCACCTCTCTGCTGCTTTCACGGGCTGGAAGACAGATGGGGCCACATCTACCTCAC	900
Db	841	CTTCACCTCTCTGCTGCTTTCACGGGCTGGAAGACAGATGGGGCCACATCTACCTCAC	900
Qy	901	GATGCTCTCTCTCAATTCGCAATCTGGGTGGCTGATCACCCTGCTCATGCTTCTCTGACTT	960
Db	901	GATGCTCTCTCTCAATTCGCAATCTGGGTGGCTGATCACCCTGCTCATGCTTCTCTGACTT	960
Qy	961	TGACCGGAGGTGGATGACACCAATCTCAGCTCCGCTTGGCTGCGCAATGGCTGGGTGTT	1020
Db	961	TGACCGGAGGTGGATGACACCAATCTCAGCTCCGCTTGGCTGCGCAATGGCTGGGTGTT	1020
Qy	1021	CTGTGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAAGCAACGAAACCCCATGGA	1080
Db	1021	CTGTGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAAGCAACGAAACCCCATGGA	1080
Qy	1081	TTATCTCTGTTGAGGATGCTTTCTGTAACCTCACTCAAGGTTTGAAGAGACAGTATGTTGGA	1140
Db	1081	TTATCTCTGTTGAGGATGCTTTCTGTAACCTCACTCAAGGTTTGAAGAGACAGTATGTTGGA	1140
Qy	1141	GAACAGAGCTTACTCTCAAGAGGAATCACTCAAGGTTTGAAGAGACAGGAGACAGCT	1200
Db	1141	GAACAGAGCTTACTCTCAAGAGGAATCACTCAAGGTTTGAAGAGACAGGAGACAGCT	1200
Qy	1201	CTATGCCCCCTATTTCCACATTTTCAAGTTCGAGAACAGCTCCCAAGAAATTTCTC	1260
Db	1201	CTATGCCCCCTATTTCCACATTTTCAAGTTCGAGAACAGCTCCCAAGAAATTTCTC	1260
Qy	1261	CATCCAGCGGCCACGCTTGGCGAGCCCTTCAAGAGCTATGAGTAAAGAAAGAGGG	1320
Db	1261	CATCCAGCGGCCACGCTTGGCGAGCCCTTCAAGAGCTATGAGTAAAGAAAGAGGG	1320
Qy	1321	CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGAGCGCGGCGGCGAGATCTAGCGGAGC	1380
Db	1321	CAGCTAACTCTGCTCTGAAGAGTGGGACAAATGAGCGCGGCGGCGAGATCTAGCGGAGC	1380
Qy	1381	TCAAAGGATGTTGGCGAAATCTTGAAGTCTTGAAGAAACTGTGAAGACATACCGGA	1440
Db	1381	TCAAAGGATGTTGGCGAAATCTTGAAGTCTTGAAGAAACTGTGAAGACATACCGGA	1440
Qy	1441	ACAGTTTGGCTCTCCCTCCAGCCTCAACCAATTTCTCCATGCTGGGCTGATGTGGCT	1500
Db	1441	ACAGTTTGGCTCTCCCTCCAGCCTCAACCAATTTCTCCATGCTGGGCTGATGTGGCT	1500
Qy	1501	AGTAAGACTCCAGTTCTTAGAGGCGCTGTAGTATTTTTTTTTTTTTTTTCTCATCCTTTGG	1560
Db	1501	AGTAAGACTCCAGTTCTTAGAGGCGCTGTAGTATTTTTTTTTTTTTTTTCTCATCCTTTGG	1560
Qy	1561	ATACTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTGAACCCCTTACTCTTTTGTGTTT	1620
Db	1561	ATACTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTGAACCCCTTACTCTTTTGTGTTT	1620
Qy	1621	GTTTTTTGAACAGGATCTTGTCTGTACCCAGGCTTGAAGTGAAGTGGGCGCATCACAG	1680
Db	1621	GTTTTTTGAACAGGATCTTGTCTGTACCCAGGCTTGAAGTGAAGTGGGCGCATCACAG	1680
Qy	1681	CCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCCATCTCCCAAGTG	1740
Db	1681	CCAGTGCAGCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCCATCTCCCAAGTG	1740
Qy	1741	CTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTCTGTTATTTT	1800
Db	1741	CTGGGATGACAGGCGTGAGCCACAGCTCCAGCCTTAGGCCCTTAATCTTCTGTTATTTT	1800

Qy	1801	CCATGACTAAAGGTCTGGTCACTCAGCTCAGCTGGCTCACAGCTCTAGGGGCTG	1860
Db	1801	CCATGACTAAAGGTCTGGTCACTCAGCTCAGCTGGCTCACAGCTCTAGGGGCTG	1860
Qy	1861	CTCCTCTAACTCAAGTGGGTTTGTGAGGCTCTGTGGCCACAGAGACCTGCATATCT	1920
Db	1861	CTCCTCTAACTCAAGTGGGTTTGTGAGGCTCTGTGGCCACAGAGACCTGCATATCT	1920
Qy	1921	GAGCAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAGGCAAC	1980
Db	1921	GAGCAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAGGCAAC	1980
Qy	1981	TTGCTGGCACCCCGCTCCCAACCTCTTGTGCTGGTAGGAGAGCTAAAGATCACCC	2040
Db	1981	TTGCTGGCACCCCGCTCCCAACCTCTTGTGCTGGTAGGAGAGCTAAAGATCACCC	2040
Qy	2041	TAAATTTACTCATCTCTTAGTGTCTCTAGTGTCTCAGCTCAGAGCTCCAGAGCA	2100
Db	2041	TAAATTTACTCATCTCTTAGTGTCTCTAGTGTCTCAGCTCAGAGCTCCAGAGCA	2100
Qy	2101	ATTCAGAGTCAACCTCTCTTCTGCACTGTCCCAACTTCTGCTCAATTCGAGATC	2160
Db	2101	ATTCAGAGTCAACCTCTCTTCTGCACTGTCCCAACTTCTGCTCAATTCGAGATC	2160
Qy	2161	TAATCTCCCTTACGCTCTGCCAGGAATTTCTTTCAGACCTCCTAGCACAGCCGGTGG	2220
Db	2161	TAATCTCCCTTACGCTCTGCCAGGAATTTCTTTCAGACCTCCTAGCACAGCCGGTGG	2220
Qy	2221	CTCCTTGTGAGGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTCT	2280
Db	2221	CTCCTTGTGAGGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGCTGATCTCT	2280
Qy	2281	CTCATCTTGCACCCCAACCTCTGTAATAGATTTACGGATTTACGGCTGCTATCTGTAA	2340
Db	2281	CTCATCTTGCACCCCAACCTCTGTAATAGATTTACGGATTTACGGCTGCTATCTGTAA	2340
Qy	2341	GTGGGATGCTCTCTTAATGGAGGAGTGTTCATTGTATATAAGTTATTACCTGAGTAT	2400
Db	2341	GTGGGATGCTCTCTTAATGGAGGAGTGTTCATTGTATATAAGTTATTACCTGAGTAT	2400
Qy	2401	GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA	2456
Db	2401	GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCAAAAAA	2456
RESULT 4			
CO894732			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1			
Rosenthal, A.D., Pillarsky, C., Dahl, E., Specht, T., Bruemendorf, T.,			
Lichtner, R., Staub, E., Roepcke, S. and Li, X.I.			
Human nucleic acid sequences expressed in pancreatic carcinomas			
Patent: EP 1471075-A 42 27-OCT-2004;			
Hinzmann, Bernd (DE); Rosenthal, Andre (DE); Pillarsky, Christian			
(DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie			
(DE)			
FEATURES			
source			
1. .6730			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
99.5%; Score 2443; DB 2; Length 6730;			

Best Local Similarity 100.0%; Pred. No. 0; Matches 2443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATAACAGCATGAAGTGCCTGGAACTGGAAATAGCGGTCTCTCCCTCGACCCCTCCCC	60						
Db	1	ATAACAGCATGAAGTGCCTGGAACTGGAAATAGCGGTCTCTCCCTCGACCCCTCCCC	60						
QY	61	TCCTTGTCCCTCTGCTCAACCCCTCGCTCGTTCCTTCCCTCCGCGAGGCGCGCTTTATA	120						
Db	61	TCCTTGTCCCTCTGCTCAACCCCTCGCTCGTTCCTTCCCTCCGCGAGGCGCGCTTTATA	120						
QY	121	ACAACTGCTCAGAGTGGGCGGGGATAGCTGCTCAAGTCTCCGCCAGCACTGAGGAG	180						
Db	121	ACAACTGCTCAGAGTGGGCGGGGATAGCTGCTCAAGTCTCCGCCAGCACTGAGGAG	180						
QY	181	CTCGCCTGTGCTCCCTCTTGCGCGGGGAGCAGCACAAGTTCAAGGCTCCGCCAGCACTGAGGAG	240						
Db	181	CTCGCCTGTGCTCCCTCTTGCGCGGGGAGCAGCACAAGTTCAAGGCTCCGCCAGCACTGAGGAG	240						
QY	241	ACTAGGGTCCAGAAATGGCTACAACAGTCCCTGATGTTGCGCAATGGGCTGAATCCAA	300						
Db	241	ACTAGGGTCCAGAAATGGCTACAACAGTCCCTGATGTTGCGCAATGGGCTGAATCCAA	300						
QY	301	GTAATAACAGACTTTGTGATAAGCTGAACTGGGCGCATCGCTAGAAACGGTGGCCAC	360						
Db	301	GTAATAACAGACTTTGTGATAAGCTGAAAGCTTGGGCGCATCGCTAGAAACGGTGGCCAC	360						
QY	361	AGCGGGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420						
Db	361	AGCGGGGTTGTGACCTCGGTGGCTTCATGCTCACTCTCCGATCCTCGTCTGCAAGGT	420						
QY	421	GCAGGACTCCAACAGCGGAAAAATGTGCTACTCAGATTTCTCTCTCTCTGCTGTT	480						
Db	421	GCAGGACTCCAACAGCGGAAAAATGTGCTACTCAGATTTCTCTCTCTCTGCTGTT	480						
QY	481	GGGCATCTTGGCTCAGCTTCGCTTCACTATCGGACTGGAGGAGCAGAGGCGCCAC	540						
Db	481	GGGCATCTTGGCTCAGCTTCGCTTCACTATCGGACTGGAGGAGCAGAGGCGCCAC	540						
QY	541	ACGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC	600						
Db	541	ACGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCTGCTGCTGCTCATGC	600						
QY	601	TGTCAGTCTGACCAAGCTCTGCGGGGAGGAAGCCCTTCCCTGTTGGTGAATCTGGG	660						
Db	601	TGTCAGTCTGACCAAGCTCTGCGGGGAGGAAGCCCTTCCCTGTTGGTGAATCTGGG	660						
QY	661	TCGCGCGTGGGCTTCAGCTAGTCCAGAGTGTATCGCTATTGAATATATTCTCTGAC	720						
Db	661	TCGCGCGTGGGCTTCAGCTAGTCCAGAGTGTATCGCTATTGAATATATTCTCTGAC	720						
QY	721	CATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCGGCTCTCGTCGCAATGAAGA	780						
Db	721	CATGAATAGGACCAACGTCATGTCTTTTCTGAGCTTTCGGCTCTCGTCGCAATGAAGA	780						
QY	781	CTTTGTCTCTGCTCACTGCTCTCTCTTTGATGGGCTGACCTTCTCTGATGCTCTC	840						
Db	781	CTTTGTCTCTGCTCACTGCTCTCTTTTCTGATGGGCTGACCTTCTCTGATGCTCTC	840						
QY	841	CTTTCACTTCTGTTGTTCTTTACGGCTGGAGAGACATGGGGCCACATCTACCTCAC	900						
Db	841	CTTTCACTTCTGTTGTTCTTTACGGCTGGAGAGACATGGGGCCACATCTACCTCAC	900						
QY	901	GATGCTCTCTCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960						
Db	901	GATGCTCTCTCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960						
QY	961	TGACCCAGTGGGATGACCACTCTCAGCTCCGCTGGCTGGCTGGCTGGCTGGCTGGT	1020						
Db	961	TGACCCAGTGGGATGACCACTCTCAGCTCCGCTGGCTGGCTGGCTGGCTGGCTGGT	1020						
QY	1021	CTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA	1080						
Db	1021	CTGTTGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAGCAACGAAACCCCATGGA	1080						
QY	1081	TTATCCTGTTGAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA	1140						
Db	1081	TTATCCTGTTGAGGATGCTTTCTGTAAACCTCAACTCGTGAAGAGAGCTATGGTGTGA	1140						
QY	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTTGAAGACAGGGGACAGCT	1200						
Db	1141	GAAACAGAGCTACTCTCAAGAGGAAATCACTCAAGGTTTTGAAGACAGGGGACAGCT	1200						
QY	1201	CTATGCCCCCTATTCCACACATTTTCAGCTGAGAAACAGCTCCCCAAAAGAAATTC	1260						
Db	1201	CTATGCCCCCTATTTCACACATTTTCAGCTGAGAAACAGCTCCCCAAAAGAAATTC	1260						
QY	1261	CATCCCAACGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG	1320						
Db	1261	CATCCCAACGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG	1320						
QY	1321	CAGCTAACTCTGCTGTAAGAGTGGGACAAATGACAGCCGGCGGAGATCTAGCGGAGC	1380						
Db	1321	CAGCTAACTCTGCTGTAAGAGTGGGACAAATGACAGCCGGCGGAGATCTAGCGGAGC	1380						
QY	1381	TCAAAGGATGTGGCGGAAATCTTGAGTCTTCTGAGAAAACCTGACAAACACTACGGGA	1440						
Db	1381	TCAAAGGATGTGGCGGAAATCTTGAGTCTTCTGAGAAAACCTGACAAACACTACGGGA	1440						
QY	1441	ACAGTTTGCTTCCCTCCAGCCTCAACCAATCTTCCATGCTGGGCTGATGTGGCT	1500						
Db	1441	ACAGTTTGCTTCCCTCCAGCCTCAACCAATCTTCCATGCTGGGCTGATGTGGCT	1500						
QY	1501	AGTAAGACTCAGTTCCTTAGAGCGCTGATATTTTTTTTTTTTTTTTGTCTCATCCTTGG	1560						
Db	1501	AGTAAGACTCAGTTCCTTAGAGCGCTGATATTTTTTTTTTTTTTTTGTCTCATCCTTGG	1560						
QY	1561	ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTGT	1620						
Db	1561	ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTCTTTTGT	1620						
QY	1621	GTTTTTTGAAAACAGGATCTGTCTGTCACCAAGCTTGAAGTGCAGTGTGGATCACAG	1680						
Db	1621	GTTTTTTGAAAACAGGATCTGTCTGTCACCAAGCTTGAAGTGCAGTGTGGATCACAG	1680						
QY	1681	CCAGTGCAGCTCCAGCACCTGTCTCAAGCAATCTCCCATCTCCATCTCCAAAGTG	1740						
Db	1681	CCAGTGCAGCTCCAGCACCTGTCTCAAGCAATCTCCCATCTCCATCTCCAAAGTG	1740						
QY	1741	CTGGGATGACAGGCGTGAAGCCCTTAAATCTTGTCTTATTTT	1800						
Db	1741	CTGGGATGACAGGCGTGAAGCCCTTAAATCTTGTCTTATTTT	1800						
QY	1801	CCATGACCTAAAGGCTGTGTCATCTGAGCTCAGCTGAGCTCAGAGCTTAGGGGCTG	1860						
Db	1801	CCATGACCTAAAGGCTGTGTCATCTGAGCTCAGCTGAGCTCAGAGCTTAGGGGCTG	1860						
QY	1861	CTCCTCTACTCAGTGGGTTTTGTGAGCTCTGTGGCCAGAGCAGACCTCATATCT	1920						
Db	1861	CTCCTCTACTCAGTGGGTTTTGTGAGCTCTGTGGCCAGAGCAGACCTCATATCT	1920						
QY	1921	GAGCAAAAATAGCAAAAGGCTCTCTCAGCCCACTGSCCTGAATCTACACTGGAAGCAAC	1980						
Db	1921	GAGCAAAAATAGCAAAAGGCTCTCTCAGCCCACTGSCCTGAATCTACACTGGAAGCAAC	1980						
QY	1981	TTGCTGGCAACCCCGCTCCCAACCCCTTGTGCTGGGTAGGAGGCTTAAAGATCACCC	2040						
Db	1981	TTGCTGGCAACCCCGCTCCCAACCCCTTGTGCTGGGTAGGAGGCTTAAAGATCACCC	2040						
QY	2041	TAAATTTACTCATCTCTAGTGCCTCAATGCGGCTCAGCAGCTCCCGAGACCA	2100						
Db	2041	TAAATTTACTCATCTCTAGTGCCTCAATGCGGCTCAGCAGCTCCCGAGACCA	2100						
QY	2101	ATTACAGGTACCCCTCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2160						
Db	2101	ATTACAGGTACCCCTCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2160						

```

QY 2161 TAATCTCCCTACGCTCTGCGAGGAATCTTTTCAGACCTCACTAGCACAAAGCCCGTTG 2220
DB 2161 TAATCTCCCTACGCTCTGCGAGGAATCTTTTCAGACCTCACTAGCACAAAGCCCGTTG 2220
QY 2221 CTCCTGTGTCAGGAGAAATTTGTAGATCAATCTCACTTCAAAATTCCTGGGGCTGATCTTCT 2280
DB 2221 CTCCTGTGTCAGGAGAAATTTGTAGATCAATCTCACTTCAAAATTCCTGGGGCTGATCTTCT 2280
QY 2281 CTATCTTGCACCCCAACCTCTGTAATAGATTTTACCGCATTTACGGCTGCATCTGTAA 2340
DB 2281 CTATCTTGTGACCCCAACCTCTGTAATAGATTTTACCGCATTTTACGGCTGCATCTGTAA 2340
QY 2341 GTGGGATGCTCTCCTTAATGGAGAGTGTTCATTTGATATAAAGTTATTACCTCTGAGTAT 2400
DB 2341 GTGGGATGCTCTCCTTAATGGAGAGTGTTCATTTGATATAAAGTTATTACCTCTGAGTAT 2400
QY 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCA 2443
DB 2401 GCAATAAAGATGTGGTGGCCACTCTTTTCATGTGGTGGCAGCA 2443

RESULT 5
BD156680
LOCUS
DEFINITION 2446 bp DNA linear PAT 17-JAN-2003
ACCESSION Primer for synthesizing full-length cDNA and use thereof.
VERSION BD156680.1 GI:27862438
KEYWORDS JP 2002191363-A/11523.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Chordata; Metazoa;
            Eukaryota;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE 1 (bases 1 to 2446)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 11523 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
            PN JP 2002191363-A/11523
            PD 09-JUL-2002
            PF 28-JUL-2000 JP 2000280990
            PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
            SAITO,
            PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
            PI KEIICHI NAGAI, TETSUJI OTSUKI
            PC
            C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ 10,
            PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
            Primer for synthesizing full-length cDNA and use thereof FH Key
            Location/Qualifiers
            FT CDS Location/Qualifiers
                1..2446
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

FEATURES
source
ORIGIN
Query Match 99.3%; Score 2439.8; DB 2; Length 2446;
Best Local Similarity 99.9%; Pred: No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATAAACAGCATGAAGTCGCGTGAACCTGGAATAGCGGTGCTCTCCCTCGACCCCTCCCC 60
DB 1 ATAAACAGCATGAAGTCGCGTGAACCTGGAATAGCGGTGCTCTCTCCCTCGACCCCTCCCC 60
QY 61 TCCTTGTCCTCTGCTCACCCCTCGCTGCTTCCCTCCCGCGAGGGCGGCTTTTATA 120
DB 61 TCCTTGTCCTCTGCTCACCCCTCGCTGCTTCCCTCCCGCGAGGGCGGCTTTTATA 120

```

```

QY 121 ACAACTGCTCAGAGTCGAGGGCGGATAGCTGTCCAAGGTCTCCCCAGCACTGAGGAG 180
DB 121 ACAACTGCTCAGAGTCGAGGGCGGATAGCTGTCCAAGGTCTCCCCAGCACTGAGGAG 180
QY 181 CTGCGCTGCTGCCCTCTTTCGCGCGGGAAGCAGCACCAAGTTTCACGGCCCAACGCTTTGGC 240
DB 181 CTGCGCTGCTGCCCTCTTTCGCGCGGGAAGCAGCACCAAGTTTCACGGCCCAACGCTTTGGC 240
QY 241 ACTAGGGTCAGAAATGGCTTACAACAGTCCCTGATGGTTGCCGAATGGCCCTGAAATCCAA 300
DB 241 ACTAGGGTCAGAAATGGCTTACAACAGTCCCTGATGGTTGCCGAATGGCCCTGAAATCCAA 300
QY 301 GTACTACAGACTTTTGTGATAAGGCTCGGGGCATCGTCTTAGAAAACGGTGGCCAC 360
DB 301 GTACTACAGACTTTTGTGATAAGGCTCGGGGCATCGTCTTAGAAAACGGTGGCCAC 360
QY 361 AGCCGGGTTGTGAACCTCGGTGGCCCTTTCATGTCTCACTCTCCGATCCTCGTCTGCAAGGT 420
DB 361 AGCCGGGTTGTGAACCTCGGTGGCCCTTTCATGTCTCACTCTCCGATCCTCGTCTGCAAGGT 420
QY 421 GCAGGACTCAACAGCGGAAATGCTGCTACTCAAGTTTCTTCTCTCTGGGTGTGTT 480
DB 421 GCAGGACTCAACAGCGGAAATGCTGCTACTCAAGTTTCTTCTCTCTGGGTGTGTT 480
QY 481 GGGCATCTTTGGCTCAACCTTCGCTTTCATCGGACTCGGACGGGAGCACAGGGCCAC 540
DB 481 GGGCATCTTTGGCTCAACCTTCGCTTTCATCGGACTCGGACGGGAGCACAGGGCCAC 540
QY 541 AGCTCTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTCTCTCTGCTGCTCATGC 600
DB 541 AGCTCTCTTCTCTTTGGGATCCTCTTTTCCATCTGCTCTCTCTGCTGCTCATGC 600
QY 601 TGTCACTGTGACCAAGCTCGTGGGGAGGAAGCCCTTTTCTCTGTTGGTGTGTTCTGGG 660
DB 601 TGTCACTGTGACCAAGCTCGTGGGGAGGAAGCCCTTTTCTCTGTTGGTGTGTTCTGGG 660
QY 661 TCTGGCCGTGGGCTTCAGGCTAGTCCAGGATGTTATCGCTATTGAATATATTGCTCTGAC 720
DB 661 TCTGGCCGTGGGCTTCAGGCTAGTCCAGGATGTTATCGCTATTGAATATATTGCTCTGAC 720
QY 721 CATGAATAGGACCAACGTCATGTCCTTTTCTGAGCTTTTCGGCTCTCGTCCGAATGAAGA 780
DB 721 CATGAATAGGACCAACGTCATGTCCTTTTCTGAGCTTTTCGGCTCTCGTCCGAATGAAGA 780
QY 781 CTTTGTCTCTGCTCAACCTAGCTCCTCTCTTGTAGTGGGCTGACCTTCTCTCATGCTCTC 840
DB 781 CTTTGTCTCTGCTCAACCTAGCTCCTCTCTTGTAGTGGGCTGACCTTCTCTCATGCTCTC 840
QY 841 CTTTCACTTCTGTGGTTCCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900
DB 841 CTTTCACTTCTGTGGTTCCTTTCACGGGCTGGAAGAGACATGGGGCCACATCTACCTCAC 900
QY 901 GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACCCTGCTCATGCTTCTGACTT 960
DB 901 GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACCCTGCTCATGCTTCTGACTT 960
QY 961 TGACCGCAGGTGGGATGACCACTCCTAGCTCCGCTTGGCTGGCCAAATGGCTGGGTGTT 1020
DB 961 TGACCGCAGGTGGGATGACCACTCCTAGCTCCGCTTGGCTGGCCAAATGGCTGGGTGTT 1020
QY 1021 CTTTGGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAGACACAAACCCCATGGA 1080
DB 1021 CTTTGGGCTTATGTTAGTCCGAGTTTGGCTGCTCACAAGACACAAACCCCATGGA 1080
QY 1081 TTATCCTGTTGAGGATGCTTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGTTGGA 1140
DB 1081 TTATCCTGTTGAGGATGCTTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGTTGGA 1140
QY 1141 GAAACAGAGCTACTCTCAAGAGAGAAATCACTCAAGGTTTTTGAAGAGACAGGGGACAGCT 1200
DB 1141 GAAACAGAGCTACTCTCAAGAGAGAAATCACTCAAGGTTTTTGAAGAGACAGGGGACAGCT 1200

```

QY 1201 CTATGCCCCCTATTCCACACATTTTCAGCTGCGAACCAGCCTCCCCAAAAGGAATTCCT 1260
Db 1201 CTATGCCCCCTATTCCACACATTTTCAGCTGCGAACCAGCCTCCCCAAAAGGAATTCCT 1260
QY 1261 CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG 1320
Db 1261 CATCCACAGGGCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGG 1320
QY 1321 CAGCTAACTCTGTCTGTAAGAGTGGACAAATGACGCGGGGCGGAGACTAGCGGGAGC 1380
Db 1321 CAGCTAACTCTGTCTGTAAGAGTGGACAAATGACGCGGGGCGGAGACTAGCGGGAGC 1380
QY 1381 TCNAAGGGATGCGGGGAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA 1440
Db 1381 TCNAAGGGATGCGGGGAAATCTTGAGTCTTCTGAGAAAACCTGTACAAGACACTACGGGA 1440
QY 1441 ACAGTTTGGCTCCCTCCCGAGCCTCAACCAATTCCTCAATGCTGGGGCTGATGGGCT 1500
Db 1441 ACAGTTTGGCTCCCTCCCGAGCCTCAACCAATTCCTCAATGCTGGGGCTGATGGGCT 1500
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
Db 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTAG 1560
QY 1561 ATACTTCTTTTAAGTGGAGTCTCAGGCACTCAAGTTTAGACCTTACTCTTTTGT 1620
Db 1561 ATACTTCTTTTAAGTGGAGTCTCAGGCACTCAAGTTTAGACCTTACTCTTTTGT 1620
QY 1621 GTTTTTTGAACAGGATCTTGCTCTGTACCCAGGCTTGAGTGCAGTGGTGCATCACAG 1680
Db 1621 GTTTTTTGAACAGGATCTTGCTCTGTACCCAGGCTTGAGTGCAGTGGTGCATCACAG 1680
QY 1681 CCAGTGCAGCTCGACCACTGTCTGTCAAGCAATCTCCATCTCCATCTCCCAAAGTG 1740
Db 1681 CCAGTGCAGCTCGACCACTGTCTGTCAAGCAATCTCCATCTCCATCTCCCAAAGTG 1740
QY 1741 CTGGGATGACAGCGGTGAGCCAGCTCCAGCCTAGGCCCTTAATCTTGCTTTATTTT 1800
Db 1741 CTGGGATGACAGCGGTGAGCCAGCTCCAGCCTAGGCCCTTAATCTTGCTTTATTTT 1800
QY 1801 CCATGACTTAAAGGCTTGGTCTCATCTGAGCTCAGCTGGGCTCACACAGCTCTAGGGCCCTG 1860
Db 1801 CCATGACTTAAAGGCTTGGTCTCATCTGAGCTCAGCTGGGCTCACACAGCTCTAGGGCCCTG 1860
QY 1861 CTCCTTAACTCACTAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
Db 1861 CTCCTTAACTCACTAGTGGGTTTGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCT 1920
QY 1921 GAGCAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAAC 1980
Db 1921 GAGCAAAATAGCAAAAGCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAAC 1980
QY 1981 TTGCTGGCACCCCGCTCCCAACCTCTTGTGCTGGGTAGGAGGCTTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCCGCTCCCAACCTCTTGTGCTGGGTAGGAGGCTTAAAGATCACCC 2040
QY 2041 TAAATTTACTCATCTCTAGTGTCTGCTCAATGCGGCTCAGAGCTCCCGACCA 2100
Db 2041 TAAATTTACTCATCTCTAGTGTCTGCTCAATGCGGCTCAGAGCTCCCGACCA 2100
QY 2101 ATTACAGGTACCCCTCTCTTGTGACTGTCCCAAACTTGTCTCAATTCAGATC 2160
Db 2101 ATTACAGGTACCCCTCTCTTGTGACTGTCCCAAACTTGTCTCAATTCAGATC 2160
QY 2161 TAAATCTCCCTTACTGCTGCAAGGAATCTTTTTCAGACCTCACTAGCAAAAGCCGGTTG 2220
Db 2161 TAAATCTCCCTTACTGCTGCAAGGAATCTTTTTCAGACCTCACTAGCAAAAGCCGGTTG 2220
QY 2221 CTCCTTGTGAGGAAATTTGTAGATCATCTTCAATTCCTGGGGCTGATATCTCT 2280
Db 2221 CTCCTTGTGAGGAAATTTGTAGATCATCTTCAATTCCTGGGGCTGATATCTCT 2280
QY 2281 CTCACTTGCACCCCAACCTCTGTAAATAGATTTCGGCAATTTACGGCTGCATCTGTAA 2340

Db 2281 CTCATCTTGCACCCCAACCTCTGTAAATAGATTTCGGCAATTTACGGCTGCATCTGTAA 2340
QY 2341 GTGGGCATGCTCTCTTAATGAGGAGTGTTCATTGTATTAATAGTTTATTCACCTGAGTAT 2400
Db 2341 GTGGGCATGCTCTCTTAATGAGGAGTGTTCATTGTATTAATAGTTTATTCACCTGAGTAT 2400
QY 2401 GCAATAAAGATGTGTGGCCACTCTTTTCATGTGTGGGAGCA 2443
Db 2401 GCAATAAAGATGTGTGGCCACTCTTTTCATGTGTGGGAGCA 2443

RESULT 6
AX877483
LOCUS AX877483 2446 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 12388 from Patent EP1074617.
ACCESSION AX877483
VERSION AX877483.1 GI:40032219
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 12388 07-FEB-2001;
Research Association for Biotechnology (JP)

FEATURES
source
1..2446
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
254..1327
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89746.1"
/db_xref="GI:40032220"

CDS
1..2446
/translation="MRTTVPDGRNGLKSKYYRLCDKABAWGIVLETVATAGVVTVA
FMULTPLLVCKVODNSNRKMLPTQFLGLVLGIFGLTFAFIIGLDGSGTPRRFFLFG
ILFSCFCLLAHAVSLTKLVGRKPLSLVILGLAVGFSLVQDVIAIEVILTMRT
NVNFEISAPRNEFDVLLTYVLFMLALTFILMSFTFCGSGTCHKRGHAIYLRNL
LSIALWVWITLLMLPDEFDRWDDTLLSALAANGWVFLAIYVSPFLLTKGRNPD
YVPEDAFCPKPLVKSYGVENRAYSQBEITQGFETGDTLYAPYSTHFLQNPPOKE
FSIPRAHAWPSPYKDYEVKKEGS"

ORIGIN

Query Match 99.3%; Score 2439.8; DB 2; Length 2446;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATAAACAGCATGAAGTCCGCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC 60
Db 1 ATAAACAGCATGAAGTCCGCTGGAACCTGGAATAGGCGTGTCTCTCCCTCGACCCCTCCCC 60
QY 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGCGCCCTTTATA 120
Db 61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCTCCCTCCGCGAGGCGCCCTTTATA 120
QY 121 ACNACTGCTCAGAGTCGAGGGGGGATAGCTGTCCAAAGTCTCCCGCAGCACTGAGGAG 180
Db 121 ACNACTGCTCAGAGTCGAGGGGGGATAGCTGTCCAAAGTCTCCCGCAGCACTGAGGAG 180
QY 181 CTCGCTGTGCTCCCTCTTGGCGCGGGAAGACAGCACAAAGTTCCACGCGCAACCGCTTTGGC 240
Db 181 CTCGCTGTGCTCCCTCTTGGCGCGGGAAGACAGCACAAAGTTCCACGCGCAACCGCTTTGGC 240
QY 241 ACTAGGGTCCAGATGGCTTACAACTGCTCCTGATGTTGCGCAATGGCTTGAATCCAA 300
Db 241 ACTAGGGTCCAGATGGCTTACAACTGCTCCTGATGTTGCGCAATGGCTTGAATCCAA 300

AK001761	LOCUS	AK001761	2446 bp	mRNA	linear	PRI 20-JAN-2006
AK001761	DEFINITION	Homo sapiens cDNA FLJ10899 fis, clone NT2RP5003506.				
AK001761	ACCESSION	AK001761				
AK001761.1	VERSION	GI:7023229				
AK001761.1	KEYWORDS	oligo capping; fis (full insert sequence).				
AK001761.1	SOURCE	Homo sapiens				
AK001761.1	ORGANISM	Homo sapiens				
AK001761.1	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.				
AK001761.1	AUTHORS	Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oba, Y., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwavanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, S., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, I. O., Nomura, Y., Iogiya, S., Komai, F., Har, R., Takeuchi, K., Arita, M., Inose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Mori, S., Momiya, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Toghiani, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isoqai, T. and Sugano, S.				
AK001761.1	TITLE	Complete sequencing and characterization of 21,243 full-length human cDNAs				
AK001761.1	JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)				
AK001761.1	PUBMED	14702039				
AK001761.1	REFERENCE	2				
AK001761.1	AUTHORS	Isoqai, T., Oca, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Kanehori, K.				
AK001761.1	TITLE	NEDO human cDNA sequencing project				
AK001761.1	JOURNAL	Unpublished				
AK001761.1	REFERENCE	3 (bases 1 to 2446)				
AK001761.1	AUTHORS	Isoqai, T. and Otsuki, T.				
AK001761.1	TITLE	Direct Submission				
AK001761.1	JOURNAL	Submitted (16-FEB-2000) Takao Isoqai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flj-cdna@ifty.com, Tel: 81-438-52-3975, Fax: 81-438-52-3986)				
AK001761.1	COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.				
AK001761.1	FEATURES	Location/Qualifiers				
AK001761.1	source	1. .2446				
AK001761.1		/organism="Homo sapiens"				
AK001761.1		/mol_type="mRNA"				
AK001761.1		/db_xref="taxon:9606"				
AK001761.1		/clone="NT2RP5003506"				
AK001761.1		/cell_line="NT2"				

AK001761.1	cell_type="teratocarcinoma"	AK001761.1	Query Match	AK001761.1	DB 5; Length 2446;
AK001761.1	/clone_lib="NT2RP5"	AK001761.1	Best Local Similarity	AK001761.1	Score 2439.8; DB 5; Length 2446;
AK001761.1	/note="cloning vector: pME18SFL3"	AK001761.1	Mismatches	AK001761.1	99.3%; Score 2439.8; DB 5; Length 2446;
AK001761.1	mRNA from NT2 neuronal precursor cells after 2-weeks	AK001761.1	Conservative	AK001761.1	99.9%; Pred. No. 0;
AK001761.1	retinoic acid (RA) induction."	AK001761.1	Matches 2441;	AK001761.1	Matches 2441; Mismatches 2; Indels 0; Gaps 0;
AK001761.1	254..1327	AK001761.1	1	AK001761.1	1
AK001761.1	/note="unnamed protein product"	AK001761.1	61	AK001761.1	61
AK001761.1	/codon_start=1	AK001761.1	61	AK001761.1	61
AK001761.1	/protein_id="BAA91890.1"	AK001761.1	121	AK001761.1	121
AK001761.1	/db_xref="GI:7023230"	AK001761.1	121	AK001761.1	121
AK001761.1	translation="MATTVPDGCNRGLKSKYVRLCDKAEAWGIVLETVATAGVVTVA	AK001761.1	181	AK001761.1	181
AK001761.1	FWLTPILLVCKVQDSNRKMLPTQELFLGLVIGLFTAFILGLDGSPTPEFLPG	AK001761.1	181	AK001761.1	181
AK001761.1	ILFSCFCSCILAHAVSLTKLVGRKPLSLVILGLAVGSLVQDVIAIYIVLTMRG	AK001761.1	241	AK001761.1	241
AK001761.1	INVSFSEISAPRNEDFVLLLYFLVFLMALTFGSGFTGWRGHAHILYLTML	AK001761.1	241	AK001761.1	241
AK001761.1	LSIAVIAWITLLMLPDEFDRWDDTILSALAAAGWVFLAYVSPFWLLTKORNDP	AK001761.1	301	AK001761.1	301
AK001761.1	YVDEAFCKPOLVKVSYGVENRAYSQEITQGFETGDTLYAPYSTHFLQNPQPE	AK001761.1	301	AK001761.1	301
AK001761.1	PSIPRAHAWPSFYKDYEVKKEGS"	AK001761.1	361	AK001761.1	361
AK001761.1		AK001761.1	361	AK001761.1	361
AK001761.1		AK001761.1	421	AK001761.1	421
AK001761.1		AK001761.1	421	AK001761.1	421
AK001761.1		AK001761.1	481	AK001761.1	481
AK001761.1		AK001761.1	481	AK001761.1	481
AK001761.1		AK001761.1	541	AK001761.1	541
AK001761.1		AK001761.1	541	AK001761.1	541
AK001761.1		AK001761.1	601	AK001761.1	601
AK001761.1		AK001761.1	601	AK001761.1	601
AK001761.1		AK001761.1	661	AK001761.1	661
AK001761.1		AK001761.1	661	AK001761.1	661
AK001761.1		AK001761.1	721	AK001761.1	721
AK001761.1		AK001761.1	721	AK001761.1	721


```
Db 721 CATGAATAGGACCAACGTCATATGCTTTCTGAGCTTTCCGCTCCTCGTCCGAATGAAGA 780
Qy CTTTGTCTCTGCTGCTCAGCTACGTCCTCTCTCTGATGCGCTGACCTTCTCTCATGTCTCTC 840
Db CTTTGTCTCTGCTGCTCAGCTACGTCCTCTCTCTGATGCGCTGACCTTCTCTCATGTCTCTC 840
Qy CTTTCACTTCTGTGGTTCTCTTACGGGCTGGAAGAGACATGGGGCCCACTACCTCAC 900
Db CTTTCACTTCTGTGGTTCTCTTACGGGCTGGAAGAGACATGGGGCCCACTACCTCAC 900
Qy GATGCTCTCTCCATTTGCGCATCTGGGTGGCTGATCAACCTGCTCATGTCTCTGACTT 960
Db GATGCTCTCTCCATTTGCGCATCTGGGTGGCTGATCAACCTGCTCATGTCTCTGACTT 960
Qy TGACCGCAGGTGGGATGACACCATCTCAGCTCGCGCTTGGCTGCGCAATGCGTGGGTGTT 1020
Db TGACCGCAGGTGGGATGACACCATCTCAGCTCGCGCTTGGCTGCGCAATGCGTGGGTGTT 1020
Qy CTTGTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAAGCAACGAAACCCCATGGA 1080
Db CTTGTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCACAAAGCAACGAAACCCCATGGA 1080
Qy TTATCTCTGTTAGGATGCTTCTGTAACCTCAACTCGTGAAGAGACTATGTGTGGA 1140
Db TTATCTCTGTTAGGATGCTTCTGTAACCTCAACTCGTGAAGAGACTATGTGTGGA 1140
Qy GAACAGAGCCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGACAGCT 1200
Db GAACAGAGCCTACTCTCAAGAGGAAATCACTCAAGGTTTGAAGAGACAGGGACAGCT 1200
Qy CTATGCCCTTATTCACACATTTTACAGTCGACAGACCGCTCCCAAGAGAAATTTCTC 1260
Db CTATGCCCTTATTCACACATTTTACAGTCGACAGACCGCTCCCAAGAGAAATTTCTC 1260
Qy CATCCCAAGGGCCCAACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG 1320
Db CATCCCAAGGGCCCAACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGGG 1320
Qy CAGCTAACTCTGTCTGAAGAGTGGGACAAATGACGCGCGGCGCAGATCTACGGGGAGC 1380
Db CAGCTAACTCTGTCTGAAGAGTGGGACAAATGACGCGCGGCGCAGATCTACGGGGAGC 1380
Qy TCAAGGGATGTGGGGGAAATCTTGAGTCTTCTGAGAAAACTGTACAGACACTACGGGA 1440
Db TCAAGGGATGTGGGGGAAATCTTGAGTCTTCTGAGAAAACTGTACAGACACTACGGGA 1440
Qy ACAGTTTGCTCTCCCTCCAGCTCAACCAAAATTTCCATGTCTGGGGCTGATGTGGCT 1500
Db ACAGTTTGCTCTCCCTCCAGCTCAACCAAAATTTCCATGTCTGGGGCTGATGTGGCT 1500
Qy AGTAAGACTCCAGTTCTTAGAGGGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTTGG 1560
Db AGTAAGACTCCAGTTCTTAGAGGGCTGTAGTATTTTTTTTTTTTTTTTGTCTCATCTTGG 1560
Qy ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTACAGCCCTTACTCTTTTGT 1620
Db ATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCAAGTTTACAGCCCTTACTCTTTTGT 1620
Qy GTTTTTTGAACAGGATCTTGTCTGTCAACCGAGCTTGTAGTGCAGTGGGCATCACAG 1680
Db GTTTTTTGAACAGGATCTTGTCTGTCAACCGAGCTTGTAGTGCAGTGGGCATCACAG 1680
Qy CCCAGTGCAGCCTCGACCATCTGTCTCAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Db CCCAGTGCAGCCTCGACCATCTGTCTCAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
Qy CTGGGATGACAGGCGTGAGCCACAGCTCCAGCTAGGCCCTTAATCTTCTGTGTTATTT 1800
Db CTGGGATGACAGGCGTGAGCCACAGCTCCAGCTAGGCCCTTAATCTTCTGTGTTATTT 1800
Qy CCATGGAATAAGGTCTGGTCACTGTAGCTCAGCTGGCTCACACAGCTCTAGGGGCTG 1860
Db CCATGGAATAAGGTCTGGTCACTGTAGCTCAGCTGGCTCACACAGCTCTAGGGGCTG 1860
```

```
Qy 1861 CTCTCTTAACCTCACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACTCGCATATCT 1920
Db 1861 CTCTCTTAACCTCACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGACTCGCATATCT 1920
Qy 1921 GAGCAAAAATAGCAAAAGCCCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAAC 1980
Db 1921 GAGCAAAAATAGCAAAAGCCCTCTCTCAGCCCACTGGCCTGAATCTACACTGGAAGCCAAC 1980
Qy 1981 TTGCTGGCACCCTCCGCTCCCAACCCCTTCTTGGCTGGTAGGAGAGCTAAAGATCACCC 2040
Db 1981 TTGCTGGCACCCTCCGCTCCCAACCCCTTCTTGGCTGGTAGGAGAGCTAAAGATCACCC 2040
Qy 2041 TAAATTTACTCATCTCTCTAGTCTGCTCACAATTGGGCTCAGCAGCTCCCCAGCACCA 2100
Db 2041 TAAATTTACTCATCTCTCTAGTCTGCTCACAATTGGGCTCAGCAGCTCCCCAGCACCA 2100
Qy 2101 ATTACAGGTCACCCCTCTCTTTCGACATGTCCCCAACTTCTGTCAATTCGAGATC 2160
Db 2101 ATTACAGGTCACCCCTCTCTTTCGACATGTCCCCAACTTCTGTCAATTCGAGATC 2160
Qy 2161 TAATCTCCCTTACGCTCTGCGAGGAATTTCTTTCAGACCTCAGTACGACAAAGCCGGTTG 2220
Db 2161 TAATCTCCCTTACGCTCTGCGAGGAATTTCTTTCAGACCTCAGTACGACAAAGCCGGTTG 2220
Qy 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGGCTGACTTCT 2280
Db 2221 CTCCTTGTGAGGAGAAATTTGTAGATCATTTCTCACTTCAAAATTCCTGGGGCTGACTTCT 2280
Qy 2281 CTCATCTTGCACCCCAACCTCTGTAATAGATTTACCGCAATTTACGGCTCATTTCTGTAA 2340
Db 2281 CTCATCTTGCACCCCAACCTCTGTAATAGATTTACCGCAATTTACGGCTCATTTCTGTAA 2340
Qy 2341 GTGGGATGCTCTCTTAATGGAGAGTGTTCATTTGATATTAATAGTTATTACCTGAGTAT 2400
Db 2341 GTGGGATGCTCTCTTAATGGAGAGTGTTCATTTGATATTAATAGTTATTACCTGAGTAT 2400
Qy 2401 GCAATTAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCA 2443
Db 2401 GCAATTAAGATGTGGTGGCCACTCTTTTCATGTTGGTGGCAGCA 2443
```

RESULT 8

AK122672
LOCUS Homo sapiens cDNA FLJ16117 fis, clone ASTRO2003632, highly smilar
DEFINITION to Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA.

AK122672
VERSION AK122672.1 GI:34527861

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, S.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yoshida, M., Horiuchi, T., Kuehno, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Ohnita, A.,
Sasaki, N., Aotake, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,

Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takenoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumaishi, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs			
TITLE	Nat. Genet. 36 (1), 40-45 (2004)		
PUBMED	14702039		
REFERENCE	2		
AUTHORS	Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M., Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T. NEDO human cDNA sequencing project		
TITLE	Unpublished		
REFERENCE	3 (bases 1 to 3057)		
AUTHORS	Isogai, T. and Yamamoto, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-JUL-2003) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flj-cdna@nifty.com, Tel: 81-438-52-3975, Fax: 81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: Reverse Proteomics Research Institute, HRI and RAB.		
FEATURES	Location/Qualifiers		
source	1. 3057 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="ASTRO2003632" /cell_type="normal astrocytes (NHA5732)" /clone_lib="ASTRO2" /notes="cloning vector: pME18SFL3 primary culture, normal astrocytes"		
ORIGIN			
Query Match	98.5%; Score 2418.4; DB 5; Length 3057;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 2419; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	1 ATAAACAGCATGAAGTCCGCTGGAACTGGATAGCGGTGCTCCCTCGACCCCTCCCC 60		
DB	638 ATAAACAGCATGAAGTCCGCTGGAACTGGATAGCGGTGCTCCCTCGACCCCTCCCC 697		
QY	61 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTTCCTCCCTCCGCGAGGGCGGCTTTATA 120		
DB	698 TCCTTGTCCCTCTGCTCACCCCTCGCTCGTTTCCTCCCTCCGCGAGGGCGGCTTTATA 757		
QY	121 ACAACTGCTCAGAGTGGGGGGGATAGCTGTCAAGGTCTCCCGCAGCACTGAGGAG 180		
DB	758 ACAACTGCTCAGAGTGGGGGGGATAGCTGTCAAGGTCTCCCGCAGCACTGAGGAG 817		
QY	181 CTCGCTGCTGCTCTTTGGCGCGGGAGCAGCACAAGTTCACGGCCAAAGCCCTTGGC 240		
DB	818 CTGCTGTGCTGCTCTTTGGCGCGGGAGCAGCACAAGTTCACGGCCAAAGCCCTTGGC 877		
QY	241 ACTAGGGTCCAGATGGCTACACAGTCCCTGATGGTTCCCGCAATGGCTGAAATCCAA 300		
DB	878 ACTAGGGTCCAGATGGCTACACAGTCCCTGATGGTTCCCGCAATGGCTGAAATCCAA 937		
QY	301 GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGCGATCGTCTTAGAACCGTGGCCAC 360		
DB	938 GTACTACAGACTTTGTGATAAGGCTGAAGCTTGGGCGATCGTCTTAGAACCGTGGCCAC 997		
QY	361 AGCCGGGGTGTGACCCCTCGGTGGCCCTTCAATGCTCACTCTCCCGATCCTCGTCTGCAAGGT 420		
DB	998 AGCCGGGGTGTGACCCCTCGGTGGCCCTTCAATGCTCACTCTCCCGATCCTCGTCTGCAAGGT 1057		
QY	421 GCAGGACTCCAAACAGCGGAAATGTCCTACTCAATGTTCTCTTCTCTCTGGGTGTGTT 480		
DB	1058 GCAGGACTCCAAACAGCGGAAATGTCCTACTCAATGTTCTCTTCTCTCTGGGTGTGTT 1117		
QY	481 GGGCATCTTTGGGCTCACCTTGCCTTTCATCATCGGACTGGAGCGGAGCACAGGGCCAC 540		
DB	1118 GGGCATCTTTGGGCTCACCTTGCCTTTCATCATCGGACTGGAGCGGAGCACAGGGCCAC 1177		
QY	541 AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTCTCTCTCTGCTGTCTGCTCATGC 600		
DB	1178 AGCTTCTTCTCTTTGGGATCCTCTTTTCCATCTCTCTCTCTGCTGTCTGCTCATGC 1237		
QY	601 TGTCACTGTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCTCTTGGTGTGTTCTGGG 660		
DB	1238 TGTCACTGTGACCAAGCTCGTCCGGGGAGGAAGCCCTTTCCTCTTGGTGTGTTCTGGG 1297		
QY	661 TCTGGCGCTGGGCTTCAGCCTAGTCCAGATGTTATCGCTATTGATATATATGTCCTGAC 720		
DB	1298 TCTGGCGCTGGGCTTCAGCCTAGTCCAGATGTTATCGCTATTGATATATATGTCCTGAC 1357		
QY	721 CATGAATAGACCAACAGCTCAATGTCTTTTCTGAGCTTTCCGCTCTCTCGCGCAATGAAGA 780		
DB	1358 CATGAATAGACCAACAGCTCAATGTCTTTTCTGAGCTTTCCGCTCTCTCGCGCAATGAAGA 1417		
QY	781 CTTTGTCTCTCTGCTCACCTAGCTCTCTTTCTGATGTTGGGCTGACCTTCTCTCATGTCCTC 840		
DB	1418 CTTTGTCTCTCTGCTCACCTAGCTCTCTTTCTGATGTTGGGCTGACCTTCTCTCATGTCCTC 1477		
QY	841 CTTCACTTCTGTGGTTCTCTTCAAGGCTGGAGAGACATGGGGCCACATCTACCTCAC 900		
DB	1478 CTTCACTTCTGTGGTTCTCTTCAAGGCTGGAGAGACATGGGGCCACATCTACCTCAC 1537		
QY	901 GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACTCTCTCATGCTTCTCTGACTT 960		
DB	1538 GATGCTCTCTCCATTTGCCATCTGGGTGGCTGGATCACTCTCTCATGCTTCTCTGACTT 1597		
QY	961 TGACCGAGGTGGGATGACACCATCTCCAGCTCCGCTGGCTGGCAATGGCTGGGTGTT 1020		
DB	1598 TGACCGAGGTGGGATGACACCATCTCCAGCTCCGCTGGCTGGCAATGGCTGGGTGTT 1657		
QY	1021 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAGAAACAGCAACCCCATGGA 1080		
DB	1658 CTTGTTGGCTTATGTTAGTCCCGAGTTTGGCTGCTCAGAAACAGCAACCCCATGGA 1717		
QY	1081 TTATCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGTTGTGA 1140		
DB	1718 TTATCTGTTGAGGATGCTTTCTGTAACCTCAACTCGTGAAGAGAGCTATGTTGTGA 1777		
QY	1141 GAACAGAGCTTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGAGACAGGGACAGCT 1200		
DB	1778 GAACAGAGCTTACTCTCAAGAGAAATCACTCAAGGTTTTTGAAGAGACAGGGACAGCT 1837		
QY	1201 CTATGCCCCCTATTTCCACATTTTTCAGCTGCAAGAACAGCTCCCAAGAAATTTCTC 1260		
DB	1838 CTATGCCCCCTATTTCCACATTTTTCAGCTGCAAGAACAGCTCCCAAGAAATTTCTC 1897		
QY	1261 CATCCCAACGGGGCCACGCTTGGCGGAGCCCTTCAAGAGCTATGAAGTAAAGAAAGGG 1320		
DB	1898 CATCCCAACGGGGCCACGCTTGGCGGAGCCCTTCAAGAGCTATGAAGTAAAGAAAGGG 1957		

QY 1321 CAGCTAACTCTGTCTGAAGATGGAGCAAAATCAGCCGGCGGCAGATCTAGCGGAGC 1380
DB |||||||
QY 1381 TCAAAGGATGTGGGGAATCTTGATCTCTTGAGAAACTGTACAAGACACTACGGGA 1440
DB |||||||
QY 2018 TCAAAGGATGTGGGGAATCTTGATCTCTTGAGAAACTGTACAAGACACTACGGGA 2077
DB |||||||
QY 1441 ACAGTTTGCTCCCTCCAGCCTCAACCAAAATCTTCCATGCTGGGCTGATGGGCT 1500
DB |||||||
QY 2078 ACAGTTTGCTCCCTCCAGCCTCAACCAAAATCTTCCATGCTGGGCTGATGGGCT 2137
DB |||||||
QY 1501 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 1560
DB |||||||
QY 2138 AGTAAGACTCCAGTTCTTAGAGCGCTGTAGTATTTTTTTTTTTTGTCTCATCTTTGG 2197
DB |||||||
QY 1561 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTTTTGT 1620
DB |||||||
QY 2198 ATACTTCTTTTAAAGTGGAGTCTCAGGCAACTCAAGTTTAGACCCCTTACTTTTGT 2257
DB |||||||
QY 1621 GTTTTTTGAACAGGATCTTGTCTGTCAACCAAGCTTGAGTCACTGGTGCATCACAG 1680
DB |||||||
QY 2258 GTTTTTTGAACAGGATCTTGTCTGTCAACCAAGCTTGAGTCACTGGTGCATCACAG 2317
DB |||||||
QY 1681 CCCAGTGCAGCCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTG 1740
DB |||||||
QY 2318 CCCAGTGCAGCCTCGACCACTGTGCTCAAGCAATCTCCCATCTCCCAAGTG 2377
DB |||||||
QY 1741 CTGGGATGACAGCGTGAAGCAGCTCCAGCTTAGGCGCTTAATCTTGTGTTATTTT 1800
DB |||||||
QY 2378 CTGGGATGACAGCGTGAAGCAGCTCCAGCTTAGGCGCTTAATCTTGTGTTATTTT 2437
DB |||||||
QY 1801 CCAATGACTAAAGGTCTGTGCTCACTGAGCTCAGCTGGCTCACAGCTCTAGGGCGCTG 1860
DB |||||||
QY 2438 CCAATGACTAAAGGTCTGTGCTCACTGAGCTCAGCTGGCTCACAGCTCTAGGGCGCTG 2497
DB |||||||
QY 1861 CTCCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGCTCATATCT 1920
DB |||||||
QY 2498 CTCCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGGCCAGAGCAGCTCATATCT 2557
DB |||||||
QY 1921 GAGCAAAATAGCAAAAGCTCTCTCAGGCCACTGGCCCTGAATCTACCTGGAAGCAAC 1980
DB |||||||
QY 2558 GAGCAAAATAGCAAAAGCTCTCTCAGGCCACTGGCCCTGAATCTACCTGGAAGCAAC 2617
DB |||||||
QY 1981 TTGCTGGCACCCCGCTCCACACCTCTTGCTGGGTAGGAGGCTAAAGATCACCC 2040
DB |||||||
QY 2618 TTGCTGGCACCCCGCTCCACACCTCTTGCTGGGTAGGAGGCTAAAGATCACCC 2677
DB |||||||
QY 2041 TAAATTTACTCATCTCTAGTGTGCTCCTCACATTTGGGCTCAGCAGCTCCCGAGCACA 2100
DB |||||||
QY 2678 TAAATTTACTCATCTCTAGTGTGCTCCTCACATTTGGGCTCAGCAGCTCCCGAGCACA 2737
DB |||||||
QY 2101 ATTACAGGTACCCCTCTCTTCTGCACTGTCCCAAACTTGTCTCAATTTCCGAGATC 2160
DB |||||||
QY 2738 ATTACAGGTACCCCTCTCTTCTGCACTGTCCCAAACTTGTCTCAATTTCCGAGATC 2797
DB |||||||
QY 2161 TAACTCCCGCTACGCTCTGCGAGGATTTCTTTCAGACCTCACTAGCAGCCCGGTTG 2220
DB |||||||
QY 2798 TAACTCCCGCTACGCTCTGCGAGGATTTCTTTCAGACCTCACTAGCAGCCCGGTTG 2857
DB |||||||
QY 2221 CTCCTTGTGAGGAAATTTGTAGATCAATCTCACTTCAAATTCCTGGGCTGATCTTCT 2280
DB |||||||
QY 2858 CTCCTTGTGAGGAAATTTGTAGATCAATCTCACTTCAAATTCCTGGGCTGATCTTCT 2917
DB |||||||
QY 2281 CTCATCTTGACCCCAACTCTGTAAATAGATTTACCGATTTTACCGCTGCAATTTGTAA 2340
DB |||||||
QY 2918 CTCATCTTGACCCCAACTCTGTAAATAGATTTACCGATTTTACCGCTGCAATTTGTAA 2977
DB |||||||
QY 2341 GTGGGATGGTCTCCTTAATGGAGGTGTTCAATTTATTAAGTTATTTCACTTGTAT 2400
DB |||||||
QY 2978 GTGGGATGGTCTCCTTAATGGAGGTGTTCAATTTATTAAGTTATTTCACTTGTAT 3037
DB |||||||
QY 2401 GCAATAAAGATGTGGTGGCC 2420

DB 3038 GCAATAAAGATGTGGTGGCC 3057
|||||
RESULT 9
AX930411
LOCUS AX930411 2302 bp DNA linear PAT 22-DEC-2003
DEFINITION Sequence 2 from Patent WO03087832.
ACCESSION AX930411
VERSION AX930411.1 GI:40312209
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Terrett, J. A.
TITLE Diagnosis of carcinoma using raigl polypeptides
JOURNAL Patent: WO 03087832-A 2 23-OCT-2003;
Oxford GlycoSciences (UK) Limited (GB)
FEATURES
source
1. .2302
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 93.7%; Score 2302; DB 2; Length 2302;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 155 CCAAGTCTCCCGGCAAGCTTGGCACTAGGAGTCCAGAAATGGCTACAAAGCTCCCTGAT 274
DB 1 CCAAGTCTCCCGGCAAGCTTGGCACTAGGAGTCCAGAAATGGCTACAAAGCTCCCTGAT 60
QY 215 ACCAAGTTCACCGCCCAACGCTTGGCACTAGGAGTCCAGAAATGGCTACAAAGCTCCCTGAT 274
DB 61 ACCAAGTTCACCGCCCAACGCTTGGCACTAGGAGTCCAGAAATGGCTACAAAGCTCCCTGAT 120
QY 275 GGTTCGCGCAATGGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTG 334
DB 121 GGTTCGCGCAATGGGCTGAAATCCAAAGTACTACAGACTTTGTGATAAGGCTGAAGCTTG 180
QY 335 GGCATCGTCTAGAACGGTGGCCACAGCGGGGTGTGACCTCGGTGGCTTCATGCTC 394
DB 181 GGCATCGTCTAGAACGGTGGCCACAGCGGGGTGTGACCTCGGTGGCTTCATGCTC 240
QY 395 ACTCTCCCGATCCTCGTCTGCAAGGTGCAAGGACTCCAAAGGCGGAAATGTGCTGCTACT 454
DB 241 ACTCTCCCGATCCTCGTCTGCAAGGTGCAAGGACTCCAAAGGCGGAAATGTGCTGCTACT 300
QY 455 CAGTTTCTCTTCTCTCGGCTGTGTTGGGCTCATCTTTGGCTCACCTTCGCTTCATCATC 514
DB 301 CAGTTTCTCTTCTCTCGGCTGTGTTGGGCTCATCTTTGGCTCACCTTCGCTTCATCATC 360
QY 515 GCACTGGAGCGGAGCAGCGGCCACACGCTTCTTCTCTTTGGGATCCTCTTTTCCATC 574
DB 361 GCACTGGAGCGGAGCAGCGGCCACACGCTTCTTCTCTTTGGGATCCTCTTTTCCATC 420
QY 575 TGCCTTCTCTGCTGCTGCTCATGTGTCAGTCTGACCAAGCTCGTCGGGGGAGGAAG 634
DB 421 TGCCTTCTCTGCTGCTGCTCATGTGTCAGTCTGACCAAGCTCGTCGGGGGAGGAAG 480
QY 635 CCCCCTTCTCTGTTGGTGAATCTGGGCTTGGCGGTGGGCTTACGCTAGTCTCAGGATGTT 694
DB 481 CCCCCTTCTCTGTTGGTGAATCTGGGCTTGGCGGTGGGCTTACGCTAGTCTCAGGATGTT 540
QY 695 ATCGCTATTGAATATATTTCTCTGACCAATAGAACCAAGCTCAATGCTCTTTCTGAG 754
DB 541 ATCGCTATTGAATATATTTCTCTGACCAATAGAACCAAGCTCAATGCTCTTTCTGAG 600
QY 755 CTTTCGCTCCTCGTGGCAATGAAGACTTTGTCTCTCTCTGCTACCTACGTCCTCTTCTTG 814

```
Db 601 CTTTCGGCTCCCTGTCGCAATGAAGACTTTGTCTCTCTGCTCACCTAGCTCCTCTTCTTG 660
Qy 815 ATGGCGCTGACCTTCTCATGTCTCTCTTCACTTCTGTGGTTCCTTACGGGCTGGAG 874
Db 661 ATGGCGCTGACCTTCTCATGTCTCTCTTCACTTCTGTGGTTCCTTACGGGCTGGAG 720
Qy 875 AGACATGGGGCCACATCTACATCTACAGATGCTCTCTCCATTTGCCATCTGGGTGCGCTGG 934
Db 721 AGACATGGGGCCACATCTACATCTACAGATGCTCTCTCCATTTGCCATCTGGGTGCGCTGG 780
Qy 935 ATCACCTGCTCATGCTTCTGACTTGTGACCCGAGTGGGATGACACCATCTCAGCTCC 994
Db 781 ATCACCTGCTCATGCTTCTGACTTGTGACCCGAGTGGGATGACACCATCTCAGCTCC 840
Qy 995 GCTTGGGTGCAATGGGCTGGTGTCTGTGGGCTTATGTTAGTCCCGAGTTTGGGCTG 1054
Db 841 GCCTTGGCTGCCAATGGGCTGGTGTCTGTGGGCTTATGTTAGTCCCGAGTTTGGGCTG 900
Qy 1055 CTCAAAAGCAACGAAACCCCATGATTAATCTGTGTGAGGATGCTTCTGTAAAACCTCAA 1114
Db 901 CTCAAAAGCAACGAAACCCCATGATTAATCTGTGTGAGGATGCTTCTGTAAAACCTCAA 960
Qy 1115 CTCGTGAAGAGACTATGCTGTGAGACAGAGCCTACTCTCAAGAGAAATCACTCAA 1174
Db 961 CTCGTGAAGAGAGCTATGCTGTGAGAAACAGAGCCTACTCTCAAGAGAAATCACTCAA 1020
Qy 1175 GCTTTTGAAGAGACAGGGACACGCTCTATGCCCCCTATTTCACACATTTTCAGCTGCGAG 1234
Db 1021 GCTTTTGAAGAGACAGGGACACGCTCTATGCCCCCTATTTCACACATTTTCAGCTGCGAG 1080
Qy 1235 AACAGCTCTCCCAAAAGGAATTCCTCATCCACGCGGCCACGCTTGGCCGAGCCCTTAC 1294
Db 1081 AACAGCTCTCCCAAAAGGAATTCCTCATCCACGCGGCCACGCTTGGCCGAGCCCTTAC 1140
Qy 1295 AAAGACTATGAAGTAAGAAAGAGGCGACTAATCTGTCTGCTGAAGTGGGACAAATGC 1354
Db 1141 AAAGACTATGAAGTAAGAAAGAGGCGACTAATCTGTCTGCTGAAGTGGGACAAATGC 1200
Qy 1355 AGCCGGCGCGCAGATCTACGGGAGCTCAAGGGATGTGGGGAATCTTGAGTCTTCTG 1414
Db 1201 AGCCGGCGCGCAGATCTACGGGAGCTCAAGGGATGTGGGGAATCTTGAGTCTTCTG 1260
Qy 1415 AGAAACTGTACAAGACACTACGGGAAAGTTTGTCTCCCTCCAGCCTCAACCAAAAT 1474
Db 1261 AGAAACTGTACAAGACACTACGGGAAAGTTTGTCTCCCTCCAGCCTCAACCAAAAT 1320
Qy 1475 CTTCCATGCTGGGCTGATGTGGCTAGTAAGACTCCAGTCTTAGAGGCGCTGTAGTAT 1534
Db 1321 CTTCCATGCTGGGCTGATGTGGCTAGTAAGACTCCAGTCTTAGAGGCGCTGTAGTAT 1380
Qy 1535 TTTTCTTTTCTCTCATCTCTTGGATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCA 1594
Db 1381 TTTTCTTTTCTCTCATCTCTTGGATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCA 1440
Qy 1595 AGTTAGACCTTACTCTTTTGTGTTTGTGTTTGAAGAGGATCTTGTCTGTCAACCGAG 1654
Db 1441 AGTTAGACCTTACTCTTTTGTGTTTGTGTTTGAAGAGGATCTTGTCTGTCAACCGAG 1500
Qy 1655 GCTTCAGTGCAGTGTGGCATCACAGCCAGTGCAGCCTCGACCACTGTGTCTCAAGCAA 1714
Db 1501 GCTTCAGTGCAGTGTGGCATCACAGCCAGTGCAGCCTCGACCACTGTGTCTCAAGCAA 1560
Qy 1715 TCCTCCCATCTCCATCTCCCAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCAGCC 1774
Db 1561 TCCTCCCATCTCCATCTCCCAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCAGCC 1620
Qy 1775 TAGGCCCTTAATCTGCTGTTATTTTCCATGACTAAAGGCTGTGCTCATCTGAGCTCAG 1834
Db 1621 TAGGCCCTTAATCTGCTGTTATTTTCCATGACTAAAGGCTGTGCTCATCTGAGCTCAG 1680
Qy 1835 CTGGCTCACAGCTCTAGGGCGCTGCTCTCTCACTCACAGTGGGTTTGTGAGGCTCT 1894
```

```
Db 1681 CTGGCTCACAGCTCTAGGGCGCTGCTCTCTAACTCACAGTGGGTTTGTGAGGCTCT 1740
Qy 1895 GTGGCCAGAGCAGACCTGCTCATATCTGAGCAAAAAATAGCAAAAGCCTCTCTCAGCCACT 1954
Db 1741 GTGGCCAGAGCAGACCTGCTCATATCTGAGCAAAAAATAGCAAAAGCCTCTCTCAGCCACT 1800
Qy 1955 GGCCTGAATCTACACTGGAAGCCAACTTGTGGGCAACCCCGCTCCCAACCTCTTTGCC 2014
Db 1801 GGCCTGAATCTACACTGGAAGCCAACTTGTGGGCAACCCCGCTCCCAACCTCTTTGCC 1860
Qy 2015 TGGGTAGGAGAGGCTAAAGATCACCTAAATTAATCTCATCTCTAGTCTGCCTCACAT 2074
Db 1861 TGGGTAGGAGAGGCTAAAGATCACCTAAATTAATCTCATCTCTAGTCTGCCTCACAT 1920
Qy 2075 TGGGCTCTCAGCAGCTCCCGCAGCACCAATTCACAGGTCAACCCCTCTCTTTTGCACCTGCC 2134
Db 1921 TGGGCTCTCAGCAGCTCCCGCAGCACCAATTCACAGGTCAACCCCTCTCTTTTGCACCTGCC 1980
Qy 2135 CCAAACTTGTCTCAATTTCCGAGATCTAATCTCCCGCTAGGCTCTGCCAGGAATTTTTC 2194
Db 1981 CCAAACTTGTCTCAATTTCCGAGATCTAATCTCCCGCTAGGCTCTGCCAGGAATTTTTC 2040
Qy 2195 AGACCTCACTAGCACAGCCCGGTTGCTCTTGTTCAGGAGAATTTGTAGATCAATCTCAC 2254
Db 2041 AGACCTCACTAGCACAGCCCGGTTGCTCTTGTTCAGGAGAATTTGTAGATCAATCTCAC 2100
Qy 2255 TTCAAATTCCTGGGCTGATCTTCTCTCATCTTGCACCCCAACCTCTCTAAATAGATTT 2314
Db 2101 TTCAAATTCCTGGGCTGATCTTCTCTCATCTTGCACCCCAACCTCTCTAAATAGATTT 2160
Qy 2315 ACCGATTTACGGCTGCATCTTGAAGTGGGATGCTCTCCTAANTGGAGAGTGTTCATT 2374
Db 2161 ACCGATTTACGGCTGCATCTTGAAGTGGGATGCTCTCCTAANTGGAGAGTGTTCATT 2220
Qy 2375 GTATAATAGTTATTTCACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTTCATGGTG 2434
Db 2221 GTATAATAGTTATTTCACCTGAGTATGCATATAAGATGTGGTGGCCACTCTTTTCATGGTG 2280
Qy 2435 GTGGCAGCAAAAAA 2456
Db 2281 GTGGCAGCAAAAAA 2302

RESULT 10
AF095448
LOCUS
DEFINITION Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA,
complete cds.
ACCESSION AF095448
VERSION AF095448.1 GI:4063889
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2302)
AUTHORS Cheng, Y. and Lotan, R.
TITLE Molecular cloning and characterization of a novel retinoic
acid-inducible gene that encodes a putative G protein-coupled
receptor
JOURNAL J. Biol. Chem. 273 (52), 35008-35015 (1998)
PUBMED 9857033
REFERENCE 2 (bases 1 to 2302)
AUTHORS Cheng, Y. and Lotan, R.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1998) Tumor Biology, The University of Texas M.
D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX
77030, USA
FEATURES
source
Location/Qualifiers
1..2302
/organism="Homo sapiens"
/mol_type="mRNA"
```



```
Db 1257 GGGACAAATGACGGCGGAGATCTAGCGGAGCTCAAGGGATGTGGCGAAATCT 1316
QY 1404 TGAGTCTTCTGAGAAAAGTGTACAGACACTACGGGAACAGTTTGGCTCCCTCCAGCCT 1463
Db 1317 TGAGTCTTCTGAGAAAAGTGTACAGACACTACGGGAACAGTTTGGCTCCCTCCAGCCT 1376
QY 1464 CAACCAAAATCTTCCATGCTGGGCTGATGTGGGCTAGTAAAGCTCCAGTCTCTAGAGG 1523
Db 1377 CAACCAAAATCTTCCATGCTGGGCTGATGTGGGCTAGTAAAGCTCCAGTCTCTAGAGG 1436
QY 1524 CGCTGAGTATTTTTTTTTTTTGTCTCATCTCTTTGGATACCTTTTAAAGTGGAGTCT 1583
Db 1437 CGCTGAGTATTTTTTTTTTTTGTCTCATCTCTTTGGATACCTTTTAAAGTGGAGTCT 1496
QY 1584 CAGGCAACTCAAGTTTAGACCTTACTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1643
Db 1497 CAGGCAACTCAAGTTTAGACCTTACTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1556
QY 1644 CTGTCAACCAAGGCTGAGTGCAGTGTGGATCACAGCCAGTGCAGCTCCGACCACTG 1703
Db 1557 CTGTCAACCAAGGCTGAGTGCAGTGTGGATCACAGCCAGTGCAGCTCCGACCACTG 1616
QY 1704 TGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTGTGGGATGACAGCGTGCAGCCAC 1763
Db 1617 TGCTCAAGCAATCTCCCATCTCCATCTCCCAAGTGTGGGATGACAGCGTGCAGCCAC 1676
QY 1764 AGCTCCAGGCTAGGCGCTTAATCTTGTGTTTATTTTCATGACTAAAGTCTGTGCTAT 1823
Db 1677 AGCTCCAGGCTAGGCGCTTAATCTTGTGTTTATTTTCCATGACTAAAGTCTGTGCTAT 1736
QY 1824 CTGAGCTCAGCGTGCAGACAGCTCAGGCGCTGCTCTCTAATCTACAGTGGGTTT 1883
Db 1737 CTGAGCTCAGCGTGCAGACAGCTCAGGCGCTGCTCTCTAATCTACAGTGGGTTT 1796
QY 1884 TGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCTGACAAATAATAGCAAAAGCTCT 1943
Db 1797 TGTGAGGCTCTGTGGCCAGAGCAGACCTGCATATCTGACAAATAATAGCAAAAGCTCT 1856
QY 1944 CTGAGCCCACTGCGCTGAATCTACCTAGGAGCAACTGTGTGGCAACCCCGCTCCCA 2003
Db 1857 CTCAGCCCACTGCGCTGAATCTACCTAGGAGCAACTGTGTGGCAACCCCGCTCCCA 1916
QY 2004 CCTTCTGCTGGGTAGGAGGCTAAAGATCACCTAAATTTACTCATCTCTAGTG 2063
Db 1917 CCTTCTGCTGGGTAGGAGGCTAAAGATCACCTAAATTTACTCATCTCTAGTG 1976
QY 2064 CTGCTCACAATTTGGGCTCAGCAGCTCCCGACCAATTTACAGGTCAACCTCTCTTC 2123
Db 1977 CTGCTCACAATTTGGGCTCAGCAGCTCCCGACCAATTTACAGGTCAACCTCTCTTC 2036
QY 2124 TTGCACTGCTCCCAAACTTGCTGTCAATTTCCGAGATCTAATCTCCCTCTACGCTGCA 2183
Db 2037 TTGCACTGCTCCCAAACTTGCTGTCAATTTCCGAGATCTAATCTCCCTCTACGCTGCA 2096
QY 2184 GGAATTTCTTTCAGACCTCACTAGCACAGCCGGTTGCTCTTGTGAGGAGAAATTTGTAG 2243
Db 2097 GGAATTTCTTTCAGACCTCACTAGCACAGCCGGTTGCTCTTGTGAGGAGAAATTTGTAG 2156
QY 2244 ATCATTTCTCAATTTCAAATTTCTGGGCTGATATCTCTCATCTTGCACCCCAACCTCTG 2303
Db 2157 ATCATTTCTCAATTTCAAATTTCTGGGCTGATATCTCTCATCTTGCACCCCAACCTCTG 2216
QY 2304 TAAATAGATTTACCGATTTACGGCTGCATTTGTAAAGTGGGATGGTCTCTTAATGGAG 2363
Db 2217 TAAATAGATTTACCGATTTACGGCTGCATTTGTAAAGTGGGATGGTCTCTTAATGGAG 2276
QY 2364 GAGTGTTCATTTAATTAAGTTATTTCACTGAGTATGAAATAAGATGTGGGCACT 2423
Db 2277 GAGTGTTCATTTAATTAAGTTATTTCACTGAGTATGAAATAAGATGTGGGCACT 2336
QY 2424 CTTTCATGTTGGTGGCAGCA 2443
```

```
2337 CTTTCATGTTGGTGGCAGCA 2356

RESULT 12
LOCUS CQ723177
DEFINITION Sequence 9111 from Patent WO02068579.
ACCESSION CQ723177
VERSION CQ723177.1 GI:42284034
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLES Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 9111 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
Location/Qualifiers
1..2290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 92.6%; Score 2274.8; DB 2; Length 2290;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2287; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 155 CCAAGGTCCTCCCCAGCACTGAGGAGCTCGCTGCTGCTTGGCGCGGGAAGCAGC 214
Db 1 CCAAGGTCCTCCCCAGCACTGAGGAGCTCGCTGCTGCTTGGCGCGGGAAGCAGC 60
QY 215 ACCAAGTTCACGGCCAAACGCTTGGCACTAGGTCAGAAATGCTACAAAGTCCCTGAT 274
Db 61 ACCAAGTTCACGGCCAAACGCTTGGCACTAGGTCAGAAATGCTACAAAGTCCCTGAT 120
QY 275 GGTGTCGCGCAATGGCTGAAATCCAAAGTACTACAGACTTTTGTGATAAGGCTGAAGCTTG 334
Db 121 GGTGTCGCGCAATGGCTGAAATCCAAAGTACTACAGACTTTTGTGATAAGGCTGAAGCTTG 180
QY 335 GGCATCGTCTTAGAAAACGGTGGCCACAGCCGGGTTGTGACCTCGGTGGCTTCATGCTC 394
Db 181 GGCATCGTCTTAGAAAACGGTGGCCACAGCCGGGTTGTGACCTCGGTGGCTTCATGCTC 240
QY 395 ACTCTCCCGATCCTCGTCTGCAAGGTGCAAGGCTCCCAACAGGCGAATAATGCTGCCTACT 454
Db 241 ACTCTCCCGATCCTCGTCTGCAAGGTGCAAGGCTCCCAACAGGCGAATAATGCTGCCTACT 300
QY 455 CAGTTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 514
Db 301 CAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 515 GGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 574
Db 361 GGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 575 TGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 634
Db 421 TGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 635 CCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 694
Db 481 CCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 695 ATCGCTATTTGAATATATTTGCTGACCAATAGGACCAACGTCATATGCTCTTTCTGAG 754
Db 541 ATCGCTATTTGAATATATTTGCTGACCAATAGGACCAACGTCATATGCTCTTTCTGAG 600
```

QY 755 CTTTCGGCTCCTCGTGGCAATGAAGACTTTGTGCTCCTCGCTCACCTAGCTCCTCTCTCTTG 814
DB 601 CTTTCGGCTCCTCGTGGCAATGAAGACTTTGTGCTCCTCGCTCACCTAGCTCCTCTCTCTTG 660
QY 815 ATGGGGCTGACCTTCTCATGTCCTCTCTCATGCTTCTGTGCTTCTTCAAGGGCTGGAG 874
DB 661 ATGGGGCTGACCTTCTCATGTCCTCTCTCATGCTTCTGTGCTTCTTCAAGGGCTGGAG 720
QY 875 AGACATGGGGCCACATCTACCTACGATGCTCTCTCCATTGCCATCTGGGTGGCCCTGG 934
DB 721 AGACATGGGGCCACATCTACCTACGATGCTCTCTCCATTGCCATCTGGGTGGCCCTGG 780
QY 935 ATCACCTGCTCATGCTTCTGATCTTTGACCCAGGTGGGATGACACCATCTCAGCTCC 994
DB 781 ATCACCTGCTCATGCTTCTGATCTTTGACCCAGGTGGGATGACACCATCTCAGCTCC 840
QY 995 GCCTTGGCTGCCAATGGCTGGGTGTTCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTG 1054
DB 841 GCCTTGGCTGCCAATGGCTGGGTGTTCTGTTGGCTTATGTTAGTCCCGAGTTTGGCTG 900
QY 1055 CTCACAAAGCAACGAAACCCCATGATTAATCTGTTGAGGATGCTTTCTGTAAACCTCAA 1114
DB 901 CTCACAAAGCAACGAAACCCCATGATTAATCTGTTGAGGATGCTTTCTGTAAACCTCAA 960
QY 1115 CTCGTGAAGAGACTATGCTGTGAGAACAGAGCTACTCTCAAGAGGAATCACTCAA 1174
DB 961 CTCGTGAAGAGACTATGCTGTGAGAACAGAGCTACTCTCAAGAGGAATCACTCAA 1020
QY 1175 GGTGTTGAAGAGACAGGGGACAGCTCTATGCCCTATTTCCACACATTTTCAGCTGCAG 1234
DB 1021 GGTGTTGAAGAGACAGGGGACAGCTCTATGCCCTATTTCCACACATTTTCAGCTGCAG 1080
QY 1235 AACCAAGCTCCCAAAAGAAATTCCTCAACGAGGCCACGCTGCGCGAGCCCTTAC 1294
DB 1081 AACCAAGCTCCCAAAAGAAATTCCTCAACGAGGCCACGCTGCGCGAGCCCTTAC 1140
QY 1295 AAAGACTAAGTAAGAAAGAGGCGAGCTAACTGTCTCTGAGAGTGGGACAAATGC 1354
DB 1141 AAAGACTAAGTAAGAAAGAGGCGAGCTAACTGTCTCTGAGAGTGGGACAAATGC 1200
QY 1355 AGCCGGCGGCGAGATCTAGCGGAGCTCAAGAGGATGTGGCGAAATCTTGAGTCTTCG 1414
DB 1201 AGCCGGCGGCGAGATCTAGCGGAGCTCAAGAGGATGTGGCGAAATCTTGAGTCTTCG 1260
QY 1415 AGAAATCTGTACAAGACACTACGGGAAAGTTTGCTCTCCCTCCAGCTCAACCAAAAT 1474
DB 1261 AGAAATCTGTACAAGACACTACGGGAAAGTTTGCTCTCCCTCCAGCTCAACCAAAAT 1320
QY 1475 CTTCCATGCTGGGCTGATGTGGGCTAGTAAGACTCCAGTCTTAGAGGCGCTAGTAT 1534
DB 1321 CTTCCATGCTGGGCTGATGTGGGCTAGTAAGACTCCAGTCTTAGAGGCGCTAGTAT 1380
QY 1535 TTTTCTTTTCTCATCTCTTGATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCA 1594
DB 1381 TTTTCTTTTCTCATCTCTTGATACTTCTTTTAAAGTGGGAGTCTCAGGCAACTCA 1440
QY 1595 AGTTAGACCCCTACTCTTTTGTGTTTGTGTTTGAACAGGATCTTGCTGTACCCAG 1654
DB 1441 AGTTAGACCCCTACTCTTTTGTGTTTGTGTTTGAACAGGATCTTGCTGTACCCAG 1500
QY 1655 GCTTGAGTGCAGTGGGCTGATCACAGCCAGTGCAGCTCGACCACTGCTGTCTCAAGCA 1714
DB 1501 GCTTGAGTGCAGTGGGCTGATCACAGCCAGTGCAGCTCGACCACTGCTGTCTCAAGCA 1560
QY 1715 TCTCTCCATCTCATCTCCCAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCAGCC 1774
DB 1561 TCTCTCCATCTCATCTCCCAAGTCTGGGATGACAGGCGTGAGCCACAGCTCCAGCC 1620
QY 1775 TAGGCCCTTAAT-CTTGCTGTTATTTTCCATGAGCTAAAGGCTGCTGATCTCAGCTAC 1833
DB 1621 TAGGCCCTTAATCTTGCTGTTATTTTCCATGAGCTAAAGGCTGCTGATCTCAGCTAC 1680
QY 1834 GCTGSGCTCACAGCTCTAGGGCCCTGCTCTTAACCTACAGTGGGTTTGTGAGGCTC 1893

DB 1681 GTGGCTCACAGCTTTAGGGCCCTGCTCTCTAACTCACAGTGGGTTTGTGAGCTC 1740
QY 1894 TGTGGCCAGAGCAGACCTGCATATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCAC 1953
DB 1741 TGTGGCCAGAGCAGACCTGCATATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCAC 1800
QY 1954 TGGCTGAATCTACACTGGAAGCAACTTGTGTGGACCCCGCTCCCAACCCCTTCTTG 2013
DB 1801 TGGCTGAATCTACACTGGAAGCAACTTGTGTGGACCCCGCTCCCAACCCCTTCTTG 1860
QY 2014 CTGGGTAGGAGGCTTAAAGATCACCTAAATTTACTCATCTCTCTAGTGTGCCCTACA 2073
DB 1861 CTGGGTAGGAGGCTTAAAGATCACCTAAATTTACTCATCTCTCTAGTGTGCCCTACA 1920
QY 2074 TTGGGCTCAGCAGCTCCCCAGCACCAATTCACAGGTCAACCCCTCTCTTTTGCATGTC 2133
DB 1921 TTGGGCTCAGCAGCTCCCCAGCACCAATTCACAGGTCAACCCCTCTCTTTTGCATGTC 1980
QY 2134 CCCAACTTGTGTCATTTCCGAGATCTAAATCTCCCCCTACGCTCTGCCAGGAATCTTT 2193
DB 1981 CCCAACTTGTGTCATTTCCGAGATCTAAATCTCCCCCTACGCTCTGCCAGGAATCTTT 2040
QY 2194 CAGACCTCACTAGCACAGCCCGTTGCTCTTGTCCAGGAAATTTGTAGATCATTTCTCA 2253
DB 2041 CAGACCTCACTAGCACAGCCCGTTGCTCTTGTCCAGGAAATTTGTAGATCATTTCTCA 2100
QY 2254 CTTCAAATTTCTGGGGCTGATATCTTCTCTCATCTTTGACACCCCAACCTCTGTAAATAGAT 2313
DB 2101 CTTCAAATTTCTGGGGCTGATATCTTCTCTCATCTTTGACACCCCAACCTCTGTAAATAGAT 2160
QY 2314 TACCGCATTTACGGCTGCATTTCTGTAAGTGGGATGCTCTCTTAATGGAGGAGTTCAT 2373
DB 2161 TACCGCATTTACGGCTGCATTTCTGTAAGTGGGATGCTCTCTTAATGGAGGAGTTCAT 2220
QY 2374 TGTATAATAAGTATTTCACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTTCATGGT 2433
DB 2221 TGTATAATAAGTATTTCACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTTCATGGT 2280
QY 2434 GGTGGCAGCA 2443
DB 2281 GGTGGCAGCA 2290

RESULT 13
AF506289
LOCUS
DEFINITION Homo sapiens orphan G protein-coupling receptor PEIG-1 mRNA, linear PRI 15-JUL-2002
ACCESSION AF506289
VERSION AF506289.1 GI:21779962
KEYWORDS complete cds.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1682 to 2040)
Cafferata,E.G., Gonzalez-Guerrico,A.M., Pivetta,O.H. and Santa-Coloma,T.A.
Identification by differential display of a mRNA specifically induced by 12-O-tetradecanoylphorbol-13-acetate (TPA) in T84 human colon carcinoma cells
Cell. Mol. Biol. 42 (5), 797-804 (1996)
8832110
2 (bases 1 to 2297)
Cafferata,E.G., Gonzalez-Guerrico,A.M., Costanzo,R., Pivetta,O.H. and Santa-Coloma,T.A.
Direct Submission
Submitted (24-APR-2002) Laboratorio de Biologia Celular y Molecular, Instituto de Investigaciones Bioquimicas Fundacion Campomar, Patricias Argentinas 435, Buenos Aires 1405, Argentina
Campomar, Location/Qualifiers

QY	1887	GAGGCTCTGTGGCCAGACAGACCTGTCATATCTGAGCAAAAATAGCAAAAGCCTCTCTC	1946
Db	1741	GAGGCTCTGTGGCCAGACAGACCTGTCATATCTGAGCAAAAATAGCAAAAGCCTCTCTC	1800
QY	1947	AGCCCACTGGCTGAATCTACACTGAAGCAACTGCTGGGACCCCGCTCCCAACCC	2006
Db	1801	AGCCCACTGGCTGAATCTACACTGAAGCAACTGCTGGGACCCCGCTCCCAACCC	1860
QY	2007	TTCTTGGCTGGGTAGGAGAGGCTAAAGATCACCCCTAAATTTACTATCTCTAGTGCTG	2066
Db	1861	TTCTTGGCTGGGTAGGATAGGCTAAAGATCACCCCTAAATTTACTATCTCTAGTGCTG	1920
QY	2067	CCTCATTGGGCTCAGAGCTGCCAGCAACCAATTCACAGTCAACCCCTCTCTCTTG	2126
Db	1921	CCTCATTGGGCTCAGAGCTGCCAGCAACCAATTCACAGTCAACCCCTCTCTCTTG	1980
QY	2127	CAGTGTCCCAACTGCTGTCAATTCAGAGATCTAATCTCCCTACGCTCTGCCAGGA	2186
Db	1981	CAGTGTCCCAACTGCTGTCAATTCAGAGATCTAATCTCCCTACGCTCTGCCAGGA	2040
QY	2187	ATTCTTTACAGCTCTACTAGCAAGCCGGTGTCTCTTGTGAGAGAAATTTAGATC	2246
Db	2041	ATTCTTTACAGCTCTACTAGCAAGCCGGTGTCTCTTGTGAGAGAAATTTAGATC	2100
QY	2247	ATTCTCACTCAAAATTCCTGGGCTGATATCTCTCATCTTTGCACCCCAACTCTGTA	2306
Db	2101	ATTCTCACTCAAAATTCCTGGGCTGATATCTCTCATCTTTGCACCCCAACTCTGTA	2160
QY	2307	ATAGATTACCGCATTTAGCGTGCATTTCTGAAGTGGCGATGGTCTCTAATGAGGAG	2366
Db	2161	ATAGATTACCGCATTTAGCGTGCATTTCTGAAGTGGCGATGGTCTCTAATGAGGAG	2220
QY	2367	TGTTCACTTATAAATGATTTCACCTCAGATGCAATAAAGATGTGTGGCCACTC	2424
Db	2221	TGTTCACTTATAAATGATTTCACCTCAGATGCAATAAAGATGTGTGGCCACTC	2278
RESULT 14	BC003665		
LOCUS	BC003665	2296 bp mRNA linear PRI 24-NOV-2004	
DEFINITION	Homo sapiens G protein-coupled receptor, family C, group 5, member A, mRNA (CDNA clone MGC:923 IMAGE:2988011), complete cds.		
ACCESSION	BC003665		
VERSION	BC003665.2	GI:33872669	
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2296)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Rah, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.G., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2296)		
AUTHORS	Director MGC Project.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	On Aug 19, 2003 this sequence version replaced gi:13177795. Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: anadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting		
FEATURES	Location/Qualifiers		
source	1..2296		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="MGC:923 IMAGE:2988011"		
	/tissue_type="Colon, adenocarcinoma"		
	/clone_lib="NIH MGC 15"		
	/lab_host="DH10B-R"		
	/notes="Vector: pOTB7"		
gene	1..2296		
	/gene="GPCR5A"		
	/note="synonyms: RAIG1, GPCR5A, RA13"		
	/db_xref="GeneID:9052"		
	/db_xref="MIM:604138"		
CDS	95..1168		
	/gene="GPCR5A"		
	/codon_start=1		
	/product="G protein-coupled receptor, family C, group 5, member A"		
	/protein_id="AAH03665.1"		
	/db_xref="GI:13177796"		
	/db_xref="GeneID:9052"		
	/db_xref="MIM:604138"		
	/translation="MATTVPDGRNGLSKYYRLCKAEAMGIVLETATAGVVTVA FMLTPLIVCKVQDSNRKMLPTQFLGLVGLIGLTFGLTAFLIGDGTGPTFFLFG ILFSCFCLLAHVSLLKLVGRKPLSLVILGLAGVSLVQDVIAIYIVLTMRRT NVNVSLSAPRNEDFVLLLTLYVLFMALTFMSSFTFGSGFTGKRGHAHLYLTWL LSIAIWAHWTLLMLPDPDRRWDITLSSALAANGHVVLLAYVSPFWLLTKORND YVPEDAFCPLVKKSYGVENRAYSQEETIQGFEETGDTLYAPSTHFLQLQPPQKE FSIPRAHAWPSPYKDYEVKKEGS"		
ORIGIN			
Query Match	92.3%;	Score 2266.4;	DB 5; Length 2296;
Best Local Similarity	99.5%;	Pred. No. 0;	
Matches 2284;	Conservative 0;	Mismatches 11;	Indels 1; Gaps 1;
QY	160	GTCTCCCCCAGCAGCTGAGGAGCTCGCTCTGCTCTTTCGCGCGGGAAGCAGCACAA	219
Db	1	GTCTCCCCCAGCAGCTGAGGAGCTCGCTCTGCTCTTTCGCGCGGGAAGCAGCACAA	60
QY	220	GTTACGGCCCAAGCCCTTGGCACTAGGTCCAGAAATGGCTACAACTCCTGATGGTTG	279
Db	61	GTTACGGCCCAAGCCCTTGGCACTAGGTCCAGAAATGGCTCCTGATGGTTG	120
QY	280	CCGCAATGCCCTGAAATCCAAGTACTACAGACTTCTGTGATAAGCTTGAAGCTTGGGCA	339

121 CGCAATGGCTGAATCCAACTACAGACTTTGTGATAAGGCTGAAGCTTTGGGCGAT 180
 340 CGTCTTAGAAAACGGTGGCCACAGCCGGGTTGTGACCTCGGTGGCTTCAATGCTCACTCT 399
 181 CGTCTTAGAAAACGGTGGCCACAGCTGGGTTGTGACCTCGGTGGCTTCAATGCTCACTCT 240
 400 CCCGATCTCTGCTGCAAGGTGAGGACTCCAAAGCGGAAAAATGCTGGCTACTCAGTT 459
 241 CCCGATCTCTGCTGCAAGGTGAGGACTCCAAAGCGGAAAAATGCTGGCTACTCAGTT 300
 460 TCTCTTCTCTGGGTGTGTGGGCACTTTTGGCCCTCACCTTCGCCCTCAATCATCGGACT 519
 301 TCTCTTCTCTGGGTGTGTGGGCACTTTTGGCCCTCACCTTCGCCCTCAATCATCGGACT 360
 520 GGAGGGGACACAGGGCCACACGCTTCTTCTCTTTTGGGATCCTCTTTTCCATCTGCTT 579
 361 GGAGGGGACACAGGGCCACACGCTTCTTCTCTTTTGGGATCCTCTTTTCCATCTGCTT 420
 580 CTCTCTGCTGCTGCTCATGCTGTGACCTGTGACCAAGCTCGTCCGGGGAGGAAGCCCT 639
 421 CTCTCTGCTGCTGCTCATGCTGTGACCTGTGACCAAGCTCGTCCGGGGAGGAAGCCCT 480
 640 TTCCCTGTTGGTAATCTGGGTGCGCGTGGCTTTCAGCCTAGTCCAGGATGTTATCGC 699
 481 TTCCCTGTTGGTAATCTGGGTGCGCGTGGCTTTCAGCCTAGTCCAGGATGTTATCGC 540
 700 TATTGAATATATGCTGACCAATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTC 759
 541 TATTGAATATATGCTGACCAATGAATAGGACCAAGCTCAATGTCTTTTCTGAGCTTTC 600
 760 CGCTCTCTCGTCCGAATGAAGACTTTTCTCTCTCTGCTCACCTAGCTCCTCTTTGATGGC 819
 601 CGCTCTCTCGTCCGAATGAAGACTTTTCTCTCTCTGCTCACCTAGCTCCTCTTTGATGGC 660
 820 GCTGACCTTCTCATGCTCTCTTCACTTCTGCTGCTTCTTCAAGGGCTGGAAGACAC 879
 661 GCTGACCTTCTCATGCTCTCTTCACTTCTGCTGCTTCTTCAAGGGCTGGAAGACAC 720
 880 TGGGGCCCACTACTACGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 939
 721 TGGGGCCCACTACTACGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 940 CTGCTCTATGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999
 781 CTGCTCTATGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 1000 GGCTGCCAATGGCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1059
 841 GGCTGCCAATGGCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 1060 AAAGCAACGAACCCCATGATTTATCTGTTGAGGATGCTTCTGTAACCTCAACTCGT 1119
 901 AAAGCAACGAACCCCATGATTTATCTGTTGAGGATGCTTCTGTAACCTCAACTCGT 960
 1120 GAAGAAGAGCTATGGTGTGGAGAACAGAGCTACTCTCAAGAGGAATCACTCAAGGTTT 1179
 961 GAAGAAGAGCTATGGTGTGGAGAACAGAGCTACTCTCAAGAGGAATCACTCAAGGTTT 1020
 1180 TGAAGAGACAGGGGACACGCTCTATGCCCTTATCCACACATTTTTCAGCTGAGAACCA 1239
 1021 TGAAGAGACAGGGGACACGCTCTATGCCCTTATCCACACATTTTTCAGCTGAGAACCA 1080
 1240 GCCTCCCCAAAGGAATTTCTCATCCCAAGGGCCAGGCTTGGCCAGGCTTCAAAAGA 1299
 1081 GCCTCCCCAAAGGAATTTCTCATCCCAAGGGCCAGGCTTGGCCAGGCTTCAAAAGA 1140
 1300 CTATGAAGTAAAGAAAGAGGAGCTAACTCTGCTGTAAGAGTGGGACAAATTCAGCGCG 1359
 1141 CTATGAAGTAAAGAAAGAGGAGCTAACTCTGCTGTAAGAGTGGGACAAATTCAGCGCG 1200
 1360 GGCGGACAGATCTAGCGGGAGCTCAAAAGGATGTGGCGAAATCTTGAGTCTTCTGAGAAA 1419
 1201 GGCGGACAGATCTAGCGGGAGCTCAAAAGGATGTGGCGAAATCTTGAGTCTTCTGAGAAA 1260

1420 ACTGTACAGACACTACGGGACAGTTTGCCTCCCTCCAGCCTCAACCAATTTCTTCC 1479
 1261 ACTGTACAGACACTACGGGACAGTTTGCCTCCCTCCAGCCTCAACCAATTTCTTCC 1320
 1480 ATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTTCTTAGAGGCGCTGTAGTA-TTTTT 1538
 1321 ATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTTCTTAGAGGCGCTGTAGTATTTTT 1380
 1539 TTTTTTTTGTCTCATCTCTTTTGGATACCTTTTAAAGTGGAGTCTCAGGCAATCTCAAGTT 1598
 1381 TTTTTTTTGGCTCATCTCTTAGGATACTTTTAAAGTGGAGTCTCAGGCAATCTCAAGTT 1440
 1599 TAGACCTTTACTCTTTTGTGTTTTTGAACAGAGTCTTCTGCTGTGCCAGGCTT 1658
 1441 TAGACCTTTACTCTTTTGTGTTTTTGAACAGAGTCTTCTGCTGTGCCAGGCTT 1500
 1659 GAGTGAGTGGTGGCATCACAGCCAGTGCAGCCTCGACCACTGTGCTCAAGCAATCCT 1718
 1501 GAGTGAGTGGTGGCATCACAGCCAGTGCAGCCTCGACCACTGTGCTCAAGCAATCCT 1560
 1719 CCATCTCCATCTCCAAAGTGTGGGATGACAGGGGTGAGGCCACAGCTCCAGCCTTAG 1778
 1561 CCATCTCCATCTCCAAAGTGTGGGATGACAGGGGTGAGGCCACAGCTCCAGCCTTAG 1620
 1779 CCCTTAATCTGCTGTATTTTTCCATGGACTAAAGGTCTGGTCACTGAGCTCAGCGTG 1838
 1621 CCCTTAATCTGCTGTATTTTTCCATGGACTAAAGGTCTGGTCACTGAGCTCAGCGTG 1680
 1839 CTCACACAGCTCTAGGGGCTGCTCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGG 1898
 1681 CTCACACAGCTCTAGGGGCTGCTCTCTCTAACTCACAGTGGGTTTTGTGAGGCTCTGTGG 1740
 1899 CCAGAGCAGACCTGCTATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCC 1958
 1741 CCAGAGCAGACCTGCTATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCACTGGCC 1800
 1959 TGAATCTACCTGGAAGCCAACTTGTGCTGCACCCCGCTCCCAACCCCTTCTTGGCTGGG 2018
 1801 TGAATCTACCTGGAAGCCAACTTGTGCTGCACCCCGCTCCCAACCCCTTCTTGGCTGGG 1860
 2019 TAGGAGAGCTAAAGATCAACCTTAAATTTACTCATCTCTAGTGTGCTGCATTCATTTGGG 2078
 1861 TAGGAGAGCTAAAGATCAACCTTAAATTTACTCATCTCTAGTGTGCTGCATTCATTTGGG 1920
 2079 CTTACAGAGCTCCCAAGCAACCAATTCAGAGTCAACCCCTCTCTTTTGCACTGTGCCCAA 2138
 1921 CTTACAGAGCTCCCAAGCAACCAATTCAGAGTCAACCCCTCTCTTTTGCACTGTGCCCAA 1980
 2139 ACTTGTGCTCAATTCGAGATCTAACTCCCTACGCTCTGCCAGGAATTTCTTTCAGAC 2198
 1981 ACTTGTGCTCAATTCGAGATCTAACTCTCCCTACGCTCTGCCAGGAATTTCTTTCAGAC 2040
 2199 CTCACCTAGCAACAGCCCGGTTGCTCTCTGTGAGGAGAAATTTGTAGATCATTTCACTTCA 2258
 2041 CTCACCTAGCAACAGCCCGGTTGCTCTCTGTGAGGAGAAATTTGTAGATCATTTCACTTCA 2100
 2259 AATTCTGGGGCTGATACCTTCTCATCTTGACCCCAACCTCTGTAAATAGATTTTACCG 2318
 2101 AATTCTGGGGCTGATACCTTCTCATCTTGACCCCAACCTCTGTAAATAGATTTTACCG 2160
 2319 CATTTACGGCTGCATTTCTGTAAGTGGCATGGTCTCTTAATGGAGAGTGTTCATTGTAT 2378
 2161 CATTTACGGCTGCATTTCTGTAAGTGGCATGGTCTCTTAATGGAGAGTGTTCATTGTAT 2220
 2379 AATAAGTATTCACTGAGTATGCAATAAAGATGTGGTGCCCACTCTTTTCACTGTGTGG 2438
 2221 AATAAGTATTCACTGAGTATGCAATAAAGATGTGGTGCCCACTCTTTTCACTGTGTGG 2280
 2439 CAGCAAAAAA 2454
 2281 AAAAAAAAAA 2296

[illegible]

Db	1443	GGGCTCAGCAGCTCCCCAGCACCATTACAGGTCAACCCCTCTCTTTGGCACTGTCCC	1502
Qy	2136	CAAACTTGCTGTCGAATCCGAGATCTAATCTCCCTTACGCTCTGCCAGGAATCTTTCA	2195
Db	1503	CAAACTTGCTGTCGAATCCGAGATCTAATCTCCCTTACGCTCTGCCAGGAATCTTTCA	1562
Qy	2196	GACCTCACTAGCACAAAGCCCGGTTGCTCTTGTGAGGAGAAATTTGTAGATCATTTCTCACT	2255
Db	1563	GACCTCACTAGCACAAAGCCCGGTTGCTCTTGTGAGGAGAAATTTGTAGATCATTTCTCACT	1622
Qy	2256	TCAAATTCCTGGGCTGATACTTCTCTCATCTTGGCACCCCAACCTCTGTAAATAGATTTA	2315
Db	1623	TCAAATTCCTGGGCTGATACTTCTCTCATCTTGGCACCCCAACCTCTGTAAATAGATTTA	1682
Qy	2316	CCGCATTTACGGCTGCATTTCTGTAAGTGGGCATGGTCTCCTAATGGAGGAGTGTTCATTG	2375
Db	1683	CCGCATTTACGGCTGCATTTCTGTAAGTGGGCATGGTCTCCTAATGGAGGAGTGTTCATTG	1742
Qy	2376	TATAATAAGTTATTACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTTCATGGTGG	2435
Db	1743	TATAATAAGTTATTACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTTCATGGTGG	1802
Qy	2436	TGGCAGCAAAAAAAAAAAAAA	2456
Db	1803	AAAAAAAAAAAAAAAAAAAAA	1823

Search completed: June 10, 2006, 18:53:49
Job time : 13886 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2006, 10:24:09 ; Search time 297 Seconds
(without alignments)
1111.887 Million cell updates/sec

Title: US-10-600-816-3

Perfect score: 1865

Sequence: 1 MATTVPDGRNGRLKSKYRL.....PRAHAWPSPKDYEVKKEGS 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1865	100.0	357	1	RA13_HUMAN	Q8nfj5 homo sapien
2	1460	78.3	357	1	Q3MHQ6_BOVIN	Q3mhq6 bos taurus
3	1402.5	75.2	356	1	RA13_MOUSE	Q8bhl4 mus musculus
4	749	40.2	188	2	Q5RK14_RAT	Q5rk14 rattus norv
5	748.5	40.1	357	2	Q6PA25_XENLA	Q6pa25 xenopus lae
6	736.5	39.5	344	1	GPC5D_MOUSE	Q9jil6 mus musculus
7	736.5	39.5	344	2	Q3UUY8_MOUSE	Q3uuy8 mus musculus
8	727.5	39.0	345	1	GPC5D_HUMAN	Q9nzd1 homo sapien
9	727.5	39.0	345	2	Q3KNV3_HUMAN	Q3knv3 homo sapien
10	605.5	32.5	432	2	Q5M7Q6_XENTR	Q5m7q6 xenopus tro
11	574	30.8	440	1	GPC5C_MOUSE	Q8k3j9 mus musculus
12	564	30.2	457	2	Q3KRC4_RAT	Q3krc4 rattus norv
13	558	29.9	441	1	GPC5B_HUMAN	Q9nc84 homo sapien
14	558	29.9	486	2	Q2NL85_HUMAN	Q2nl85 homo sapien
15	557	29.9	452	2	Q5JGM2_HUMAN	Q5jgm2 homo sapien
16	547.5	29.4	468	2	Q2YDGO_BOVIN	Q2ydg0 bos taurus
17	545	29.2	408	2	Q5M7K5_XENTR	Q5m7k5 xenopus tro
18	523	28.0	410	1	GPC5B_MOUSE	Q923z0 mus musculus
19	508	27.2	403	1	GPC5B_HUMAN	Q9nzh0 homo sapien
20	498	26.7	403	2	Q5R9B5_PONPY	Q5r9b5 pongo pygma
21	491	26.3	403	2	Q5R6C3_PONPY	Q5r6c3 pongo pygma
22	458.5	24.6	353	2	Q4SGJ6_TETNG	Q4sgj6 tetraodon n
23	439	23.5	393	2	Q4RJM3_TETNG	Q4rjm3 tetraodon n
24	413.5	22.2	349	2	Q4R812_TETNG	Q4r812 tetraodon n
25	344.5	18.5	537	2	Q4RZM7_TETNG	Q4rzx7 tetraodon n
26	269	14.4	262	2	Q9JMF0_MOUSE	Q9jmf0 mus musculus
27	166	8.9	105	2	Q9BSP0_HUMAN	Q9bsp0 homo sapien
28	152.5	8.2	896	2	Q8MSJ2_DROME	Q8msj2 drosophila
29	151.5	8.1	1404	2	Q20073_CAEEL	Q20073 caenorhabdi
30	151	8.1	896	1	BOSS_DROME	P22815 drosophila
31	151	8.1	917	2	Q24265_DROME	Q24265 drosophila

RESULT 1

ID	RA13_HUMAN	STANDARD;	PRT;	357 AA.
AC	Q8NFJ5; O95357;			
DT	13-APR-2004, integrated into UniProtKB/Swiss-Prot.			
DT	13-APR-2004, sequence version 2.			
DT	07-FEB-2006, entry version 28.			
DE	Retinoic acid-induced protein 3 (G-protein coupled receptor family C group 5 member A) (Retinoic acid-induced gene 1 protein) (RAIG-1) (Orphan G-protein coupling receptor PEIG-1).			
DE	Name=GPC5A; Synonyms=RA13, RAIG1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [MRNA].			
RC	TISSUE=Lung;			
RX	MEDLINE=99074278; PubMed=9857033; DOI=10.1074/jbc.273.52.35008;			
RA	Cheng Y., Iotani R.;			
RT	"Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G protein-coupled receptor.";			
RL	J. Biol. Chem. 273:35008-35015(1998).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [MRNA].			
RC	TISSUE=Colon carcinoma;			
RA	Cafferata E.G., Gonzalez-Guerrico A.M., Costanzo R., Pivetta O.H.,			
RA	Santa-Coloma T.A.;			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Oiyama M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,			
RA	Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,			
RA	Sudo H., Hosori T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,			
RA	Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,			
RA	Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,			
RA	Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y., Iahida S.,			
RA	Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Nomura J.,			
RA	Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,			
RA	Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,			
RA	Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,			
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,			
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,			
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,			

Q24738 drosophila
Q6pat5 mus musculus
Q9qys2 mus musculus
Q624e7 caenorhabdi
Q46p7 tetraodon n
Q49xb9 tetraodon n
P31422 rattus norv
Q14832 homo sapien
Q5ral3 pongo pygma
Q4t846 tetraodon n
Q86y93 homo sapien
Q14416 homo sapien
Q59ge5 homo sapien
Q73555 fugu rubrip

RA Kawabata A., Hikiji T., Kobatake N., Inegaki H., Ikema Y., Okamoto S.,
RA Oktani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RL cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE:Colon;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP TISSUE SPECIFICITY
RX MEDLINE=20247251; PubMed=10783259; DOI=10.1006/geno.2000.6164;
RA Brauner-Oberne H., Krosgaard-Larsen P.;
RT "Sequence and expression pattern of a novel human orphan G-protein-
RT coupled receptor, GPRC5B, a family C receptor with a short amino-
RT terminal domain.";
RL Genomics 65:121-128(2000).
RN [6]
RP TISSUE SPECIFICITY
RX PubMed=10945465; DOI=10.1006/geno.2000.6226;
RA Robbins M.J., Michalovich D., Hill J., Calver A.R., Medhurst A.D.,
RA Gloger I., Sims M.A., Middlemiss D.N., Pangalos M.N.;
RT "Molecular cloning and characterization of two novel retinoic acid-
RT inducible orphan G-protein-coupled receptors (GPRC5B and GPRC5C).";
RL Genomics 67:8-18(2000).
CC -!- FUNCTION: Unknown. This G-protein coupled receptor could be
CC involved in modulating differentiation and maintaining homeostasis
CC of epithelial cells. The comparable expression level in fetal lung
CC and kidney with adult tissues suggests a possible role in
CC embryonic development and maturation of these organs. This
CC retinoic acid-inducible GPCR provide evidence for a possible
CC interaction between retinoid and G-protein signaling pathways.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in the
CC plasma membrane and perinuclear vesicles.
CC -!- TISSUE SPECIFICITY: Expressed at high level in fetal and adult
CC lung tissues. Constitutively expressed in fetal kidney and adult
CC placenta, kidney, prostate, testis, ovary, small intestine, colon,
CC stomach, and spinal chord at low to moderate levels. Not
CC detectable in fetal heart, brain, and liver and adult heart,
CC brain, liver, skeletal muscle, pancreas, spleen, thymus, and
CC peripheral leukocytes. According to Ref.5: expressed at low but
CC detectable level in pancreas and heart.
CC -!- INDUCTION: By all-trans retinoic acid (ATRA).
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

DR EMBL; AF095448; AAC98506.1; -; mRNA.
DR EMBL; AF506289; AA077594.1; -; mRNA.
DR EMBL; AK001761; BAA91890.1; -; mRNA.
DR EMBL; BC003665; AA03665.1; -; mRNA.
DR Ensembl; ENSG00000013588; Homo sapiens.
DR HGNC; HGNC:9836; GPRC5A.
DR MIM; 604138; Gene.
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0004930; F: G-protein coupled receptor activity; TAS.
DR GO; GO:0007165; P: signal transduction; TAS.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; FALSE_NEG.
DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR PROSITE; PS02059; G_PROTEIN_RECEP_F3_4; FALSE_NEG.
KW G-protein coupled receptor; Glycoprotein; Membrane; Polymorphism;
KW Receptor; Transducer; Transmembrane
FT CHAIN 1 357 Retinoic acid-induced protein 3.
FT 1 357 /FTID=PRO_0000206895.
FT TOPO_DOM 1 33 Extracellular (Potential).
FT TRANSMEM 34 54 1 (Potential).
FT TOPO_DOM 55 68 Cytoplasmic (Potential).
FT TRANSMEM 69 89 2 (Potential).
FT TOPO_DOM 90 97 Extracellular (Potential).
FT TRANSMEM 98 118 3 (Potential).
FT TOPO_DOM 119 129 Cytoplasmic (Potential).
FT TRANSMEM 130 150 4 (Potential).
FT TOPO_DOM 151 176 Extracellular (Potential).
FT TRANSMEM 177 197 5 (Potential).
FT TOPO_DOM 198 212 Cytoplasmic (Potential).
FT TRANSMEM 213 233 6 (Potential).
FT TOPO_DOM 234 247 Extracellular (Potential).
FT TRANSMEM 248 268 7 (Potential).
FT TOPO_DOM 269 357 Cytoplasmic (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT VARIANT 118 118 S>G (in dbSNP:850932).
FT 292 /FTID=VAR_018296.
FT CONFLICT 292 S->G (in Ref. 2).
FT SEQUENCE 357 AA; 40251 MW; 7BB524BF6F307E5 CRC64;
SQ
Query Match 100.0%; Score 1865; DB 1; Length 357;
Best Local Similarity 100.0%; Pred. No. 9.2e-138; Indels 0; Gaps 0;
Matches 357; Conservative 0; Mismatches 0;
QY 1 MATTVPDCRCNGLSKYYRLCDKAEAWGIVLETATAGVWTSVAFMLTLPILVCKVQDSN 60
Db |||||
1 MATTVPDCRCNGLSKYYRLCDKAEAWGIVLETATAGVWTSVAFMLTLPILVCKVQDSN 60
QY 61 RKKMLPTQFLFLGLVIGIFGLTFAPFIIGLDGSTGTRFRLFGILFISCFSCLLAHAVSLT 120
Db |||||
61 RKKMLPTQFLFLGLVIGIFGLTFAPFIIGLDGSTGTRFRLFGILFISCFSCLLAHAVSLT 120
QY 121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMRNTNVNVSFELSAPRNEDFVLL 180
Db |||||
121 KLVGRKPLSLVILGLAVGFSLVQDVIAEIVLTMRNTNVNVSFELSAPRNEDFVLL 180
QY 181 LTYVLFMALTFMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAIVTLLMLPDFDRW 240
Db |||||
181 LTYVLFMALTFMSSFTFCGSFTGKRGHAIYLTMLLSIAIWAIVTLLMLPDFDRW 240
QY 241 DDTILSSALAANGWVFLAYSPFWLLTKQPNMDYVEDAFCKPOLVKKSYGVENRAY 300
Db |||||
241 DDTILSSALAANGWVFLAYSPFWLLTKQPNMDYVEDAFCKPOLVKKSYGVENRAY 300
QY 301 SOEELTQCFEETGDTLYAPYSTHFOLOQPPKQFSPRAHAWPSFYKDYVKKEGS 357
Db |||||
301 SOEELTQCFEETGDTLYAPYSTHFOLOQPPKQFSPRAHAWPSFYKDYVKKEGS 357
RESULT 2
Q3MHQ6_BOVIN PRELIMINARY; PRT; 357 AA.
ID Q3MHQ6_BOVIN

```
AC Q3MH06;
DT 25-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2005, sequence version 1.
DT 07-MAR-2006, entry version 6.
DE Hypothetical protein MGCI27231.
GN Name=MGCI27231;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Crossbred x Angus; TISSUE=ileum;
RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
RA Tanguchi M., Wang Z., Yu J., Frange C., Schreiber K., Shenmen C.,
RA Wagner L., Balu M., Barbazuk S., Barber S., Babakaiff R., Beland J.,
RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
RA Mateuo C., Mayo M., Santos R.R., Scott J., Tsai M., Wong D.,
RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC105147; RA105148.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008067; F:metabotropic glutamate, GABA-B-like recepto. . ; IEA.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm 3; 1.
KW Hypothetical protein.
SQ SEQUENCE 357 AA; 39917 MW; CB5C6B59CFEF5B6C CRC64;

Query Match 78.3%; Score 1460; DB 2; Length 357;
Best Local Similarity 76.5%; Pred. No. 5.1e-106;
Matches 273; Conservative 38; Mismatches 46; Indels 0; Gaps 0;

QY 1 MATTPDGRNGLSKRYRLCDKAEAWGVLETVATAGVVTVAFMLTPIILVCKVQDSN 60
DB 1 MSTVAPAGCRSLDSRYRLCDKAAWGVLEALAGVAVAFMITLLVLICKVQDSN 60

QY 61 RRKMLPTQFLFLGLVIGLFAFIIGDGTGTPTRFELFGLFSLFCFCLLAHAVSLT 120
DB 61 KRKLLPTQFLFLGLVIGLFAFIITLDGGTPTPTFELFGLFALCFCLLVHAFNLT 120

QY 121 KLVGRKPLSLVILGLAVFSLQVDIAIEIVILVTMNTNVTNVSSELSAPRNEDFVLL 180
DB 121 KLVGRQPLSLVILGLALGSLVQDIIAIEYVLTMTNRTNVTNVSSELSPPRNEDFVML 180

QY 181 LTVVLFLMALTFMSSFTFCGSGFTGWRHGAHYLTMLLSIAIWAIVITLLMLPDRRW 240
DB 181 LTVVLFLMALTFELTASFQSGFTAWKRHGAHYLTMLLSIAIWAIVITLLVPTLSPQM 240

QY 241 DDTILSSALAANGWFLAYVSPFEWLLTKQRNPMDPYVEDAFCKPQLVKVSGYGVENRAY 300
DB 241 DDTILSSALVANGWFLAYIAPEFHLITRQQNPMDPYVEDAFQCPQFMKQYGVVNRAY 300

QY 301 SQEETITQGEFTGDTLYAPYSTHFQIQNQPPQKPEFSPRAHAWPSYKDYVKKSGS 357
DB 301 SQEIIITQGEETGDTLYAPYSTHFQIQNRDSAKDFSIPRVQTRVSPYSDEYGRKDVS 357

RESULT 3
RA13 MOUSE
AC Q8BH14; Q8K1G5;
DT 13-APR-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Retinoic acid-induced protein 3 (G-protein coupled receptor family C
DE group 5 member A) (Retinoic acid-induced gene 1 protein) (RAIG-1).
GN Name=Gprc5a; Synonyms=Ra13, Ra1gl;
OS Mus musculus (Mouse).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129/SV; TISSUE=Embryonic stem cell;
RX PubMed=14706456; DOI=10.1016/S0888-7543(03)00237-4;
RA Tao Q., Cheng Y., Clifford J., Lotan R.;
RT "Characterization of the murine orphan G-protein-coupled receptor gene
RL Ra13 and its regulation by retinoic acid.";
Genomics 83:270-280(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=PVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywiecki M.I., Skalska U., Smalilus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Unknown. This G-protein coupled receptor could be
CC involved in modulating differentiation and maintaining homeostasis
CC of epithelial cells. This retinoic acid-inducible GPCR provide
CC evidence for a possible interaction between retinoid and G-protein
CC signaling pathways.
CC -!- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed predominantly in normal fetal and
CC adult lung. Almost undetectable or expressed at very low levels in
CC other tissues.
CC -!- INDUCTION: By all-trans retinoic acid (ATRA).
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY079518; AAL87526.1; ALT_INIT; Genomic DNA.
DR EMBL; AY079517; AAL87526.1; JOINED; Genomic DNA.
DR EMBL; BC034063; AAH34063.1; -; mRNA.
DR EMBL; BC036173; AAH36173.1; -; mRNA.
DR EMBL; BC039217; AAH39217.1; -; mRNA.
DR Ensemble; ENSMUSG00000046733; Mus musculus.
DR MGI; MGI:1891250; Gprc5a.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm 3; 1.
DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE NEG.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; FALSE NEG.
DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE NEG.
DR PROSITE; PS00959; G_PROTEIN_RECEP_F3_4; FALSE NEG.
KW G-protein coupled receptor; Glycoprotein; Membrane; Receptor;
KW Transducer; Transmembrane.
CHAIN 1 356 Retinoic acid-induced protein 3.
/FTID=PRO_000206896.
TOPO_DOM 1 35 Extracellular (Potential).
TRANSMEM 36 56 1 (Potential).
TOPO_DOM 57 68 Cytoplasmic (Potential).
TRANSMEM 69 89 2 (Potential).
TOPO_DOM 90 101 Extracellular (Potential).
```

```
FT TRANSMEM 102 122 3 (Potential).
FT TOPO DOM 123 131 Cytoplasmic (Potential).
FT TRANSMEM 132 152 4 (Potential).
FT TOPO DOM 153 178 Extracellular (Potential).
FT TRANSMEM 179 199 5 (Potential).
FT TOPO DOM 200 214 Cytoplasmic (Potential).
FT TRANSMEM 215 235 6 (Potential).
FT TOPO DOM 236 244 Extracellular (Potential).
FT TRANSMEM 245 265 7 (Potential).
FT TOPO DOM 266 356 Cytoplasmic (Potential).
FT CARBOHYD 160 160 N-linked (GlcNAc...) (Potential).
FT CONFLICT 136 196 T -> A (in Ref. 1).
SQ SEQUENCE 356 AA; 40101 MW; 08B632F228381C01 CRC64;

Query Match 75.2%; Score 1402.5; DB 1; Length 356;
Best Local Similarity 76.1%; Pred. No. 1.6e-101;
Matches 270; Conservative 26; Mismatches 56; Indels 3; Gaps 2;

QY 3 TTVPDGRNGLKSKYRILCDKAEANGIVLETATAGVTSVAFMLTLPILVCKVQDSNRR 62
DB 5 TTAPSGCRSDLSRYHRLCDLAEGWGIALETTLAAVGAVATACMFALVFLICKVQDSNKR 64

QY 63 KMLPTQFLGLVGLGFTFAFIIGLDGSGTGTFFELFGILFSCFCLLAHAVSLTKL 122
DB 65 KMLPAQFLGLVGLGFTFAFIIGLDGSGTGTFFELFGILFSCFCLLAHAVSLTKL 124

QY 123 VGRGRPLSLVILGLAVGFSLVQDVIAIEYIVLTWNRTNVNVFSELSAPRNEEDFVLLIT 182
DB 125 VGRGRPLSLVILGLAVGFSLVQDVIAIEYIVLTWNRTNVNVFSELSAPRNEEDFVLLIT 184

QY 183 YVFLMALTFMLMSFTFCGSGTGWKRHGAHYLTMLLSIAIWMITLMLPDFRRWDD 242
DB 185 YVFLVLMVLTFTFSLVFCGSGTGWKRHGAHYLTMLLSIAIWMITLMLPDFRRWDD 244

QY 243 TILSSALANGVFLAYVSPFWLLTKORNPMDYPVEDAFCKPOLVKSGYVENRAYSQ 302
DB 245 TILSTALVANGVFLAYVSPFWLLTKORNPMDYPVEDAFCKPOLVKSGYVENRAYSQ 304

QY 303 BEITQGTETGDTLYAPYSTHFLQNLQNPQPKQFSPRAHAMPSPYKDYEVKKEGS 357
DB 305 BEITQGL-EMGDTLYAPYSTHFLQNLQNH--QKDFSIPRAQAPASPYNDYEGRKGS 356

RESULT 4
QSRK14 RAT PRELIMINARY; PRT; 188 AA.
AC QSRK14
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2006, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Ra13 protein (Fragment).
GN Name=Ra13;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
```



```
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield F.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX Klein S., Strausberg R., to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC060480; AAH60480.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008067; F:metabotropic glutamate, GABA-B-like recepto. .; IEA.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm 3; 1.
DR PRINTS; PR00248; GPCR_MGR.
DR PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
SQ SEQUENCE 357 AA; 40250 MW; 42CFE07FA09E5E4 CRC64;

Query Match 40.1%; Score 748.5; DB 2; Length 357;
Best Local Similarity 46.5%; Pred. No. 2.9e-50;
Matches 160; Conservative 62; Mismatches 105; Indels 17; Gaps 7;

QY 20 LCKAEAWGIVLETAGVTVSVARMLTLPILVCKVQDSNRKMLPTQPLFLGLVGLIF 79
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12 LCTDAAGVIVLETLAAAGVIVFSIIILALLIMPRISDPKRAVSPVQLIFLIGTFGIF 71

QY 80 GLTFAFIIGDGTPTREFLFGILFISCSLLAHAVSITKLVRGKPLSLVILGLAV 139
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 GLTFAFIVLETCPTREFLFGVLFVFAICPSCLLAHASKLVRLVRGGHGCWMMMLMAL 131

QY 140 GFSLVQDVIAIEYIVLTMRN--TNVNVFS-ELSAPRRNEDFVLLTYVFLMALTFMSS 196
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 FLPLVQVVIALLIVLGLVRSCTPTVFNGATVYHQLNQDFVLLIYVFLMAITFLVSL 191

QY 197 FPCGSGFTGWRGHAHYLTMLSLIAIWMITLLMLPDFD-----RRWDDTILSSALAA 251
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 ISLCGPKCYKRGHAHYIVTMFFSIGIWIWAWICMLLRGNADLNQGNRTWDDPVLISALVA 251

QY 252 NGWVELLAYVSPFELLTKORNPWDYVEDAFCKPOLVKSYGVENRAYSOEITOGFEE 311
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
252 NGWVELMVMVPELCMLTRCQS--DRQDCVCQTPHLLRTIGVDRVFTHENTNOG-QD 308

QY 312 TGTLYAPYSTH----FQLQNPPQKEFSIPRAHAWPSPYKDYE 351
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
309 SGRC--SPVSSRODATIAMEDLEQKXDFSIPRQPRQNPYMQQ 350

RESULT 6
GPC5D MOUSE
ID GPC5D MOUSE STANDARD; PRT; 344 AA.
AC Q9JIL6; Q8CJ10; Q8CJ11;
DT 13-APR-2004, integrated into UniProtKB/Swiss-Prot.
DT 13-APR-2004, sequence version 2.
DT 07-FEB-2006, entry version 25.
DE G-protein coupled receptor family C group 5 member D.
GN Names=Gprc5d;
```

```
FT TOPO DOM 189 204 Cytoplasmic (Potential).
FT TRANSMEM 205 225 6 (Potential).
FT TOPO DOM 226 239 Extracellular (Potential).
FT TRANSMEM 240 260 7 (Potential).
FT TOPO DOM 261 344 Cytoplasmic (Potential).
FT VARSPLIC 299 300 AR -> EC (in isoform 2).
FT VARSPLIC 301 344 /FTID=VSP 010009.
FT VARSPLIC 344 344 Missing (in isoform 2).
FT CONFLICT 24 24 /FTID=VSP 010010.
FT CONFLICT 234 235 I -> C (in Ref. 3).
FT CONFLICT 235 235 OP -> HA (in Ref. 3).
SQ SEQUENCE 344 AA; 39056 MW; CBEC1B6A3F683764 CRC64;

Query Match 39.5%; Score 736.5; DB 1; Length 344;
Best Local Similarity 45.4%; Pred. No. 2.5e-49;
Matches 153; Conservative 57; Mismatches 116; Indels 11; Gaps 6;

QY 17 YRLCDKAEANGIVLETATAGVTSVAFMLTLPLVCKVQDSNRKMLPTQFLFLLGVL 76
DB 12 YFLFCNDEGPMVAIVLESIAVIGIVVITLLLAFLFMKRVQDCSQNNVLTQFLFLAVL 71
QY 77 GFGLTFAFIIGLDSGTGPTFFLFGILFISCFCLLAHVASLTKLVRGRKPLSLVILG 136
DB 72 GLFGLTFAFIIGLDSGTGPTFFLFGILFISCFCLLAHVASLTKLVRGRVSVFCWTTILF 131
QY 137 LAVGSLVQDVIAEYIVLTNRTNVNVPSELSPARNEDFVLLTYVFLMALTFMLSS 196
DB 132 IAGVSLLTQITIAIEYVTLITRGT-LMFEHTPTQLNDFVCLLYVFLMALTFVFSK 189
QY 197 FTFCSFTGWEKGAHIVLTMLSTAIWVATLML--PDPRD--WDDTILSALAAN 252
DB 150 ATFCCGPNWKGHRLIFATVLVSIIVVWVSMLLRGNPQLQRPQPHDDVAICITGLVTN 249
QY 253 GWVFLVAVSPFWMLLTKQRNPMYDVEDAFCKPOLVKVSYGVENRAYSQEITQGFEE 312
DB 250 ANVFLIIYIIPLSILYRSR--QECPTQGNVCQVPVQSRFMDTQETRADSDGAB- 307
QY 313 GDTLVAPYSTHQLQNPQKQFSTIPRAHAMPSPYKD 349
DB 308 -DVALTAYGTPTQLQSDPSPREYLIPSATL--SPOQD 341

RESULT 7
ID Q3UUY8_MOUSE PRELIMINARY; PRT; 344 AA.
AC Q3UUY8;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE 6 days neonate skin cDNA, RIKEN full-length enriched library,
DE clone:A030001P22 product:G protein-coupled receptor, family C, group
DE 5, member D, full insert sequence.
GN Name=Gprc5d;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Skin;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Skin;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
```

```
RA Ambesi-Impombato A., Apweiler R., Auraliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Clougherty V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furum P., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnicki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasaki Y., Kelson J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nissson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J.C., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarogov E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Kayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kohjima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Skin;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grammond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
```


FT	VARSP LIC	299	345	ARDSGAEEDVALTSGTPIQPTQVDPDTECEIPQAKLSPPQ QDAGCV -> GTFLDGSRSRELLQEKQKNHVG (in isoform 3). /FTid=VSP_010008.
FT	VARSP LIC	299	300	AR -> DC (in isoform 2). /FTid=VSP_010006.
FT	VARSP LIC	301	345	Missing (in isoform 2). /FTid=VSP_010007.
FT	VARIANT	18	18	A -> D (in dbSNP:3741822). /FTid=VAR_018297.
SQ	SEQUENCE	345 AA;	38791 MW;	14809EA4C5E9F9B49 CRC64;
Query Match 39.0%; Score 727.5; DB 1; Length 345;				
Best Local Similarity 45.7%; Pred. No. 1.2e-48;				
Matches 154; Conservative 55; Mismatches 117; Indels 11; Gaps 6				
QY	17	YYRLCDKAEANGIVLEIVATAGVVTSVAFMLTLT	PLVCKVODSNRKMPLPTQFLFLGLVL 76	
DB	12	YFLCCDAEGPWGIILESALIGIVVTILLAFILMKRI	QDCSOMNVLP TOLLFLLSVL 71	
QY	77	GIFGLTFAFIIIGDSTGPTFRFFLFGILFSCFSC	LLAHAVSLTKLVRGRKPLSLVILG 136	
DB	72	GLFGLAFAFIIELNQQTAPRYFLFGLVFLCFC	SLLAHANLVKLVRCGVCSFSTTILC 131	
QY	137	LAGFSIVQDVIALEYIVLTNNRTNNVFSLSAPR	NEDFVLLTYVFLMALTFMLSS 196	
DB	132	IAIGCSLLQIIIAEYVTLIMTRG--MMFVNMT	PCQLNVDFVLLVYVFLMALTFVFSK 189	
QY	197	FTFCGSTGKRGCAHIYLTMLLSIAIWAIVTILML	--PDFDR--RWDDETILSSALAN 252	
DB	190	ATFCGPCENKQGRILFITVLFISIIWVMISML	LRGNPQFQRPQMDPVPVICALVTN 249	
QY	253	GWVFLLAYSPFEWLLTKQRNPMDPYVEDAF	CKPQLVKKSYGVENRAYSQEITOGFEET 312	
DB	250	AWVFLLYIVPELCILYRSCR-QBCPLQGNAC	PYTAIOHVSQVENQELSRARDSGAE-- 307	
QY	313	GDTLYAPYSTHFQIQNQPPQKFSIPRAHAW	SPYKD 349	
DB	308	-DVALTSYGTPIQPTQVDPDTECEIPQAKL-	-SPQD 341	
RESULT 9				
ID	Q3KNV3 HUMAN	PRELIMINARY;	PRT;	345 AA.
AC	Q3KNV3;			
DT	08-NOV-2005,	integrated into UniProtKB/TrEMBL.		
DT	08-NOV-2005,	sequence version 1.		
DE	07-FEB-2006,	entry version 4.		
DE	G protein-coupled receptor,	family C, group 5, member D.		
OS	Name=GPRC5D;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa;	Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria;	Euarchontoglires; Primates; Catarrhini; Homnidae;		
OC	Homo.			
OX	NCBI_TaxID=9606;			
ON	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=PCR rescued clones;			
RX	MEDLINE=22386257; PubMed=12477932;	DOI=10.1073/pnas.242603899;		
RA	Klausner R.L., Feingold E.A., Grouse L.H.,	Derge J.G.,		
RA	Strausberg R.D., Collins F.S., Wagner L.,	Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H.,	Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T.,	Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A.,	Rubin G.M., Hong L.,		
RA	Scapleton M., Soares M.B., Bonaldo M.F.,	Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S.,	Carninci P., Frange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J.,	Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J.,	Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S.,	Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J.,	Lu X., Gibbs R.A.,		
RA	Faney J., Helton E., Kettenman M., Madan A.,	Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C.,	Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.C.,	Dickson M.C.,		

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC107078; AAI07079.1; -; mRNA.
DR EMBL; BC107077; AAI07078.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008067; F:metabotropic glutamate, GABA-B-like recepto. .; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005118; F:sevenless binding; IEA.
DR GO; GO:0007601; P:visual perception; IEA.
DR InterPro; IPR002956; Bridge_of_7less.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm 3; 1.
DR PRINTS; PR01223; BRIDBOF7LESS.
KW Receptor.
SQ SEQUENCE 345 AA; 38791 MW; 14B09E4C5E9F9B49 CRC64;

Query Match 39.0%; Score 727.5; DB 2; Length 345;
Best Local Similarity 45.7%; Pred. No. 1.2e-48;
Matches 154; Conservative 55; Mismatches 117; Indels 11; Gaps 6;

QY 17 YRLCDKAEANGIVLTATAGVTVTSVAFMLTLPILVCKVQDSNRRMLPTQLFLGLVL 76
DB 12 YFLLCDAECPGWGILSLAIGLVTVVILLAFLLMRKIQDSQWNVLPTQLFLLSVL 71
QY 77 GFGLTFATFIIGDGTPTFRFLFGILFSCILAHAVSLTKLVGRKPLSLAVILG 136
DB 72 GUGFLAFATIELNQOTAPVRFLFGLFALCFSCILAHASNLKLVGRVCSFVSTTILC 131
QY 137 LAVGSLVQDVIAIEIVLTMTNRTNVNVSFELSAPRNEFDVLLTYVLFMLATFLMS 196
DB 132 IAIGCSLLQIIITATEVTVTLIMTRG--MMFVNMTPCQLNVDFVLLVYVLFMLATFFVSK 189
QY 197 FTFCGSGTGWKRGAHYITMLLSIAIIVANITLLML--PDPR--RWDITLSSALAN 252
DB 190 ATFCGCPENWKQGRILFIFTVLFSIIIVVWISMLLRGNPQFORQPDQDPPVVCIALVTN 249
QY 253 GWVFLAYVSPFVLLTKQRPMDYVEDAFCKPQLVKVSYGVENRAYSQEEITQGFET 312
DB 250 ANWFLLLYVPELCILYRSCR--QECPLQGNACPVTAHQSHFQVENOELSRADSDGAER- 307
QY 313 GDTLXAPYSTHFQIQNPQPKFESIPRAHAWSPYKD 349
DB 308 -DVALTSYGTPIQPVTDPTQECFIQAKL--SPQD 341

RESULT 10
QSM706 XENTR PRELIMINARY; PRT; 432 AA.
AC QSM7067
DT 01-FEB-2005, integrated into UniProtKB/TREMBL.
DT 01-FEB-2005, sequence version 1.
DE 07-FEB-2006, entry version 9.
DE Hypothetical LOC496814.
GN Names:LOC496814;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC089508; AAH89508.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008067; F:metabotropic glutamate, GABA-B-like recepto. .; IEA.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm 3; 1.
KW Hypothetical protein.
SQ SEQUENCE 432 AA; 48246 MW; 2DAB4751E7203433 CRC64;

Query Match 32.5%; Score 605.5; DB 2; Length 432;
Best Local Similarity 37.6%; Pred. No. 5.8e-39;
Matches 143; Conservative 66; Mismatches 120; Indels 51; Gaps 15;

QY 3 TTVPDGRNGLSKYVRLCDKAEANGIVLTATAGVTVTSVAFMLTLPILVCK---VQDS 59
DB 30 TAVPSGCGAGINSLYYLCDLNAAAGIVLQAITAGIVST--FILTI-VLVAFTPIQNK 86
QY 60 NRRKMLPTQFLGLVGLGIFGLTFAPIIGLDGSGTPTFRFLFGILFSCILAHAVSL 119
DB 87 KKSLLGTQVFLGLTGLGIFGLVDFVDMVKDPATCASRRFLGLLFAICFACLWVHGVS 146
QY 120 TKLVGRKPLSLVILGLAVGSLVQDVIAIEIVLTMTNRTNVNVSFELSAPR-----RN 174
DB 147 NYLIRNTQPSGWIWIFGLAVCLSLVEAIINTEWLIITIVKT-----ANPPADPCLIGN 199
QY 175 EDFVLLTYVLFMLATFLMSSFTFCGSGTGWKRGAHYITMLLSIAIIVANITLLMLP 234
DB 200 ADFVMAIYVFMFLIAGFITAPALCGRYGHWKKAIFILLTLFUSIWIWVIMVYVG 259
QY 235 DFD-----RWDDTILSSALAANGVFLAYVSPFVLLTK-----QNPMDYPVE--- 280
DB 260 NVQVGNPLYWDDPTLAIALVSGNAFIFFYIIPETISQLTKPTLEQTFESEP--YPIRGVG 317
QY 281 -DAFCPKQLVKVSYGVENRAYSQEEIT-----QGFE-ETGDTLYAPYSTHFQIQNP 329
DB 318 YETILKEQSTQRMV-VENKAFSMDPEPPSAKPKVPSVSGYNGQTKGVQYQ--TEMALMNK 374
QY 330 P-PQKEFS--IPRAHAWSP 346
DB 375 TLPSYDVIVIPRATANPPP 394

THIS PAGE IS BLANK

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:18:44 ; Search time 1145.2 Seconds
(without alignments)
2345.260 Million cell updates/sec

Title: US-10-600-816-30

Perfect score: 42

Sequence: 1 gccacgcttgccgagcccc.....acaaagactatgaagtaaa 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vl.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	42	100.0	BD233463	BD233463 Human pro
2	42	100.0	BD209699	BD209699 Compositi
3	42	100.0	AR341505	AR341505 Sequence
4	42	100.0	AX078375	AX078375 Sequence
5	42	100.0	BD233473	BD233473 Human pro
6	42	100.0	AK172760	AK172760 Homo sapi
7	42	100.0	CQ723177	CQ723177 Sequence
8	42	100.0	BC003665	BC003665 Homo sapi
9	42	100.0	AF506289	AF506289 Homo sapi
10	42	100.0	AF506289	AF506289 Homo sapi
11	42	100.0	AF095448	AF095448 Homo sapi
12	42	100.0	BD156680	BD156680 Primer fo
13	42	100.0	AX877483	AX877483 Sequence
14	42	100.0	AK001761	AK001761 Homo sapi
15	42	100.0	CQ981495	CQ981495 Sequence
16	42	100.0	DD210040	DD210040 Methods o
17	42	100.0	AX549168	AX549168 Sequence
18	42	100.0	AK122672	AK122672 Homo sapi

19	42	100.0	3371	2	AX188348	AX188348 Sequence	
20	42	100.0	6730	2	CQ894732	CQ894732 Sequence	
C	21	42	100.0	161577	5	AC007688	Homo sapi
22	33.6	80.0	680	2	BD150346	Primer fo	
23	33.6	80.0	680	2	AX870284	Sequence	
24	26	61.9	194709	12	AC132640	Rattus no	
25	26	61.9	232234	12	AC108634	Rattus no	
C	26	61.9	240216	12	AC107096	Rattus no	
27	24.6	58.6	235468	6	AC105586	Rattus no	
28	24.6	58.6	235631	12	AC126155	Rattus no	
29	24.4	58.1	189833	12	AC116169	Homo sapi	
30	24	57.1	229303	12	AC157012	Bos tauru	
31	23.6	56.2	161566	5	AC008060	Homo sapi	
C	32	23	54.8	135629	12	CT573208	Danio res
C	33	23	54.8	169003	12	CT009616	Danio res
34	23	54.8	185146	6	AC159967	Mus muscu	
35	23	54.8	230474	6	AC102564	Mus muscu	
36	22.8	54.3	1500	2	AR630602	Sequence	
37	22.8	54.3	1500	10	AF112297	Japanese	
38	22.8	54.3	1500	10	AF112298	Japanese	
39	22.8	54.3	1500	10	AY243826	Japanese	
40	22.8	54.3	1500	10	AY243843	Japanese	
41	22.8	54.3	1500	10	JEU44974	Japanese en	
42	22.8	54.3	1500	10	JEU70389	Japanese en	
43	22.8	54.3	1500	10	JEU70413	Japanese en	
44	22.8	54.3	1500	10	JEU70419	Japanese en	
45	22.8	54.3	1500	10	JEU70420	Japanese en	
46	22.8	54.3	1500	10	JEU70420	Japanese en	
47	22.8	54.3	1896	10	AY184214	Japanese	
48	22.8	54.3	2069	14	BC105147	Bos tauru	
49	22.8	54.3	2379	10	S75726	Japanese en	
50	22.8	54.3	2436	10	JEU03696	Japanese en	
51	22.8	54.3	4320	10	JEU03696	Japanese en	
52	22.8	54.3	4512	2	AR003118	Sequence	
53	22.8	54.3	4512	2	AR003211	Sequence	
54	22.8	54.3	4992	10	JEU03696	Japanese en	
55	22.8	54.3	4992	2	I08692	Sequence 2	
56	22.8	54.3	6536	2	BD064747	Expressio	
57	22.8	54.3	7486	2	BD064684	Preparati	
58	22.8	54.3	7486	2	BD064685	Preparati	
59	22.8	54.3	7486	2	BD064746	Expressio	
60	22.8	54.3	10976	10	JEU03696	Japanese en	
C	61	22.8	54.3	10976	10	JEU03696	Japanese en
C	62	22.8	54.3	141695	6	AC121911	Mus muscu
C	63	22.8	54.3	145213	6	AC142410	Mus muscu
C	64	22.8	54.3	247690	12	AC095186	Rattus no
65	22.8	54.3	249669	12	AC137204	Rattus no	
66	22.8	54.3	298216	12	AC006875	Caenorhab	
67	22.6	53.8	161575	12	AC141041	Rattus no	
68	22.6	53.8	173071	12	AC166012	Oryctolag	
C	69	22.6	53.8	188622	12	AC091441	Homo sapi
C	70	22.4	53.3	1857	13	DQ211932	Toxoplasma
C	71	22.4	53.3	11959	13	AC024754	Caenorhab
C	72	22.4	53.3	34483	5	AC135368	Homo sapi
C	73	22.4	53.3	110000	15	GF522870	21 c
C	74	22.4	53.3	115880	15	PFU344068	Pseudomon
C	75	22.4	53.3	136857	5	AC068600	Homo sapi
76	22.4	53.3	159453	12	AC006732	Caenorhab	
77	22.4	53.3	176355	5	AC025518	Homo sapi	
C	78	22.4	53.3	181914	12	AC024640	Homo sapi
C	79	22.4	53.3	186870	12	AC116170	Homo sapi
C	80	22.2	52.9	584	2	AX400400	Sequence
81	22.2	52.9	69203	12	AC104541	Mus muscu	
82	22.2	52.9	73878	12	AC011600	3 o	
83	22.2	52.9	110000	12	AC011600	2	
C	84	22.2	52.9	156766	5	AC079597	Homo sapi
C	85	22.2	52.9	205274	12	AC128161	Rattus no
C	86	22.2	52.9	218866	11	AC171018	Gallus ga
C	87	22.2	52.9	226291	12	AC096068	Rattus no
88	22.2	52.9	247515	12	AC126713	Rattus no	
C	89	22	52.4	167735	6	AC122818	Mus muscu
C	90	22	52.4	190256	6	AC121302	Mus muscu
C	91	22	52.4	191710	6	AC122489	Mus muscu

C	92	22	52.4	203589	6	AC084383	AC084383 Mus muscu	165	21.4	51.0	2356	2	CQ574206	CQ574206 Sequence	
C	93	22	52.4	209153	12	AC092090	AC092090 Canis fam	166	21.4	51.0	2368	13	AF132165	AF132165 Drosophil	
C	94	22	52.4	249159	12	AC111880	AC111880 Rattus no	167	21.4	51.0	3579	13	AY904042	AY904042 Trypanoso	
C	95	22	52.4	255568	12	AC127823	AC127823 Rattus no	168	21.4	51.0	4880	10	SPV1GAG	X58484 Simian foam	
C	96	21.8	51.9	428	11	AY844358	AY844358 Acris cre	169	21.4	51.0	13972	10	SPVGENOME	X54482 Simian foam	
C	97	21.8	51.9	428	11	AY844359	AY844359 Acris gry	170	21.4	51.0	32182	13	DQ011274	DQ011274 Oikopleur	
C	98	21.8	51.9	428	11	AY844389	AY844389 Hyia arbo	171	21.4	51.0	53394	5	AC013276	AC013276 Homo sapi	
C	99	21.8	51.9	428	11	AY844439	AY844439 Hyia aren	c	172	51.0	58558	12	AC024305	AC024305 Homo sapi	
C	100	21.8	51.9	428	11	AY844440	AY844440 Hyia japo	c	173	51.0	110000	4	AF008213	AF008213 110	
C	101	21.8	51.9	428	11	AY844443	AY844443 Hyia mixe	c	174	51.0	119895	15	AP004310	AP004310 Synchocy	
C	102	21.8	51.9	428	11	AY844462	AY844462 Hyia equi	c	175	51.0	120469	12	AC158875	AC158875 Bos tauru	
C	103	21.8	51.9	428	11	AY844465	AY844465 Hyia vers	c	176	51.0	125648	12	AC158857	AC158857 Bos tauru	
C	104	21.8	51.9	428	11	AY844466	AY844466 Hyia walk	c	177	51.0	143668	4	AP004268	AP004268 Oryza sat	
C	105	21.8	51.9	428	11	AY844504	AY844504 Pseudacri	c	178	51.0	145342	12	AC179493	AC179493 Strongylo	
C	106	21.8	51.9	428	11	AY844519	AY844519 Scinax fu	c	179	51.0	147826	12	AC141723	AC141723 Apis mell	
C	107	21.8	51.9	734	7	BV630333	BV630333 S217P6119	c	180	51.0	177982	12	AC155545	AC155545 Zea mays	
C	108	21.8	51.9	1339	2	AX646217	AX646217 Sequence	c	181	51.0	182523	12	AC182208	AC182208 Bos tauru	
C	109	21.8	51.9	1339	5	AB065829	AB065829 Homo sapi	c	182	51.0	193237	12	AC179031	AC179031 Strongylo	
C	110	21.8	51.9	1509	11	AY583347	AY583347 Chlogloss	c	183	51.0	211717	12	AC175161	AC175161 Bos tauru	
C	111	21.8	51.9	1512	11	AY571662	AY571662 Hyia meri	c	184	51.0	222378	12	AC015862	AC015862 Homo sapi	
C	112	21.8	51.9	1512	11	AY583339	AY583339 Hyia meri	c	185	51.0	237488	12	AC095201	AC095201 Rattus no	
C	113	21.8	51.9	21670	5	AF508041	AF508041 Homo sapi	c	186	51.0	243969	11	AC145956	AC145956 Gallus ga	
C	114	21.8	51.9	21748	2	CQ595097	CQ595097 Sequence	c	187	51.0	245036	12	AC094925	AC094925 Rattus no	
C	115	21.8	51.9	36628	5	AF037222	AF037222 Human DNA	c	188	51.0	258110	12	AC095639	AC095639 Rattus no	
C	116	21.8	51.9	55704	5	CT009610	CT009610 Human DNA	c	189	51.0	306349	12	AC145065	AC145065 Pan trogl	
C	117	21.8	51.9	95961	12	AC017329	AC017329 Drosophil	c	190	21.2	50.5	286	10	AB029292	AB029292 Japanese
C	118	21.8	51.9	100598	5	AC006271	AC006271 Homo sapi	c	191	21.2	50.5	286	10	AB029293	AB029293 Japanese
C	119	21.8	51.9	143136	13	AC023677	AC023677 Drosophil	c	192	21.2	50.5	286	10	AB029294	AB029294 Japanese
C	120	21.8	51.9	146466	5	AL353637	AL353637 Human DNA	c	193	21.2	50.5	286	10	AB029295	AB029295 Japanese
C	121	21.8	51.9	156415	5	AC011464	AC011464 Homo sapi	c	194	21.2	50.5	286	10	AB029296	AB029296 Japanese
C	122	21.8	51.9	165617	5	CNS0000K	AL049840 Human chr	c	195	21.2	50.5	286	10	AB029300	AB029300 Japanese
C	123	21.8	51.9	167531	5	AL158832	AL158832 Human DNA	c	196	21.2	50.5	286	10	AB029301	AB029301 Japanese
C	124	21.8	51.9	175378	12	AC113179	AC113179 Mus muscu	c	197	21.2	50.5	618	10	AY501273	AY501273 HIV-1 iso
C	125	21.8	51.9	175679	13	AC010025	AC010025 Drosophil	c	198	21.2	50.5	1482	10	AF036915	AF036915 Japanese
C	126	21.8	51.9	178159	12	AC173237	AC173237 Bos tauru	c	199	21.2	50.5	1500	10	AB028248	AB028248 Japanese
C	127	21.8	51.9	181194	6	AL844531	AL844531 Mouse DNA	c	200	21.2	50.5	1500	10	AB028249	AB028249 Japanese
C	128	21.8	51.9	193461	5	AC023403	AC023403 Homo sapi	c	201	21.2	50.5	1500	10	AB028250	AB028250 Japanese
C	129	21.8	51.9	206834	12	AC135066	AC135066 Homo sapi	c	202	21.2	50.5	1500	10	AB028251	AB028251 Japanese
C	130	21.8	51.9	207858	12	AC127132	AC127132 Rattus no	c	203	21.2	50.5	1500	10	AB028271	AB028271 Japanese
C	131	21.8	51.9	229892	12	AC096895	AC096895 Rattus no	c	204	21.2	50.5	1500	10	AY029207	AY029207 Japanese
C	132	21.8	51.9	248669	12	AC127097	AC127097 Rattus no	c	205	21.2	50.5	1500	10	AY243827	AY243827 Japanese
C	133	21.8	51.9	253387	12	AC133379	AC133379 Rattus no	c	206	21.2	50.5	1500	10	AY243831	AY243831 Japanese
C	134	21.8	51.9	297381	13	AE003540	AE003540 Drosophil	c	207	21.2	50.5	1500	10	AY243833	AY243833 Japanese
C	135	21.6	51.4	452	2	BD040609	BD040609 Sequence	c	208	21.2	50.5	1500	10	AY243837	AY243837 Japanese
C	136	21.6	51.4	452	2	AR740870	AR740870 Sequence	c	209	21.2	50.5	1500	10	AY243839	AY243839 Japanese
C	137	21.6	51.4	452	2	AX905076	AX905076 Sequence	c	210	21.2	50.5	1500	10	AY555762	AY555762 Japanese
C	138	21.6	51.4	65851	12	AC172133	AC172133 Bos tauru	c	211	21.2	50.5	1500	10	AY555764	AY555764 Japanese
C	139	21.6	51.4	70000	2	AR215862	AR215862 Sequence	c	212	21.2	50.5	1500	10	AY555765	AY555765 Japanese
C	140	21.6	51.4	73317	5	AL522921	AL522921 Homo sapi	c	213	21.2	50.5	1500	10	AY555766	AY555766 Japanese
C	141	21.6	51.4	95375	5	HS228A9	AL022322 Human DNA	c	214	21.2	50.5	1500	10	AY555767	AY555767 Japanese
C	142	21.6	51.4	122590	4	AC148346	AC148346 Medicago	c	215	21.2	50.5	1500	10	JEU44958	JEU44958 Japanese en
C	143	21.6	51.4	127054	12	AC163249	AC163249 Loxodonta	c	216	21.2	50.5	1500	10	JEU44960	JEU44960 Japanese en
C	144	21.6	51.4	127270	5	AC036144	AC036144 Homo sapi	c	217	21.2	50.5	1500	10	JEU44964	JEU44964 Japanese en
C	145	21.6	51.4	129081	12	AC163247	AC163247 Loxodonta	c	218	21.2	50.5	1500	10	JEU44966	JEU44966 Japanese en
C	146	21.6	51.4	144824	6	CR936848	CR936848 Mouse DNA	c	219	21.2	50.5	1500	10	JEU44967	JEU44967 Japanese en
C	147	21.6	51.4	149103	12	AC128053	AC128053 Rattus no	c	220	21.2	50.5	1500	10	JEU44968	JEU44968 Japanese en
C	148	21.6	51.4	168829	5	AC009941	AC009941 Homo sapi	c	221	21.2	50.5	1500	10	JEU44969	JEU44969 Japanese en
C	149	21.6	51.4	179591	5	AC093902	AC093902 Homo sapi	c	222	21.2	50.5	1500	10	JEU44973	JEU44973 Japanese en
C	150	21.6	51.4	183224	12	AC048363	AC048363 Homo sapi	c	223	21.2	50.5	1500	10	JEU70395	JEU70395 Japanese en
C	151	21.6	51.4	190469	6	AC109172	AC109172 Mus muscu	c	224	21.2	50.5	1500	10	JEU70396	JEU70396 Japanese en
C	152	21.6	51.4	207878	12	AC013566	AC013566 Homo sapi	c	225	21.2	50.5	1500	10	JEU70399	JEU70399 Japanese en
C	153	21.6	51.4	213518	12	AC165932	AC165932 Bos tauru	c	226	21.2	50.5	1500	10	JEU70400	JEU70400 Japanese en
C	154	21.6	51.4	221115	12	AC111437	AC111437 Rattus no	c	227	21.2	50.5	1500	10	JEU70405	JEU70405 Japanese en
C	155	21.6	51.4	226967	6	AL627425	AL627425 Mouse DNA	c	228	21.2	50.5	1500	10	JEU70411	JEU70411 Japanese en
C	156	21.6	51.4	233589	6	AL683824	AL683824 Mouse DNA	c	229	21.2	50.5	1500	10	JEU70412	JEU70412 Japanese en
C	157	21.6	51.4	237899	6	AC102622	AC102622 Mus muscu	c	230	21.2	50.5	1500	10	JEU70421	JEU70421 Japanese en
C	158	21.6	51.4	240349	12	AC096133	AC096133 Rattus no	c	231	21.2	50.5	1500	10	JEVE61	JEVE61 Japanese en
C	159	21.4	51.0	407	2	AX916898	AX916898 Sequence	c	232	21.2	50.5	1500	10	S47265	S47265 E protein [
C	160	21.4	51.0	407	2	BD052431	BD052431 Sequence	c	233	21.2	50.5	2435	10	JEU21054	JEU21054 Japanese en
C	161	21.4	51.0	407	2	AR752692	AR752692 Sequence	c	234	21.2	50.5	2436	10	JEU03691	JEU03691 Japanese en
C	162	21.4	51.0	428	11	AY844413	AY844413 Hyia cycl	c	235	21.2	50.5	2436	10	JEU03692	JEU03692 Japanese en
C	163	21.4	51.0	428	11	AY844438	AY844438 Hyia neph	c	236	21.2	50.5	2436	10	JEU03694	JEU03694 Japanese en
C	164	21.4	51.0	428	11	AY844487	AY844487 Osteopiliu	c	237	21.2	50.5	2436	10	JEU21057	JEU21057 Japanese en

238	21.2	50.5	4406	10	AY027863	AY027863 Japanese	C 311	21	50.0	692	2	CQ411716	CQ411716 Sequence
239	21.2	50.5	5853	13	JBANATAT1	M20871 T. brucei va	C 312	21	50.0	758	7	BV563202	BV563202 qcz74g09.
240	21.2	50.5	6254	4	AE622846	AJ622826 Emericea en	C 313	21	50.0	1008	5	AY453396	AY453396 Homo sapi
241	21.2	50.5	10951	10	JEVLINGCG	L78128 Japanese en	314	21	50.0	1008	5	BT007150	BT007150 Homo sapi
242	21.2	50.5	10970	10	AF254452	AF254452 Japanese	315	21	50.0	1008	5	BT020135	BT020135 Homo sapi
243	21.2	50.5	10970	10	AF254453	AF254453 Japanese	316	21	50.0	1008	8	AY888499	AY888499 Synthetic
244	21.2	50.5	10970	10	AY303791	AY303791 Japanese	317	21	50.0	1008	8	AY890300	AY890300 Synthetic
245	21.2	50.5	10970	10	AY303792	AY303792 Japanese	318	21	50.0	1008	8	AY892776	AY892776 Synthetic
246	21.2	50.5	10976	10	AF014160	AF014160 Japanese	319	21	50.0	1008	8	BT008192	BT008192 Synthetic
247	21.2	50.5	10976	10	AF014161	AF014161 Japanese	320	21	50.0	1096	2	CQ413562	CQ413562 Sequence
248	21.2	50.5	10976	10	AF098735	AF098735 Japanese	321	21	50.0	1386	2	CQ413739	CQ413739 Sequence
249	21.2	50.5	10976	10	AF098736	AF098736 Japanese	322	21	50.0	1580	5	AB028741	AB028741 Homo sapi
250	21.2	50.5	10976	10	AF098737	AF098737 Japanese	323	21	50.0	1586	5	BC014613	BC014613 Homo sapi
251	21.2	50.5	10976	10	AY508812	AY508812 Japanese	324	21	50.0	1587	14	BC102576	BC102576 Bos tauru
252	21.2	50.5	10976	10	AY508813	AY508813 Japanese	325	21	50.0	1599	6	BC034063	BC034063 Mus muscu
253	21.2	50.5	10977	10	AB196923	AB196923 Japanese	326	21	50.0	1648	2	AX405638	AX405638 Sequence
254	21.2	50.5	10977	10	AB196924	AB196924 Japanese	327	21	50.0	1936	6	BC036174	BC036174 Mus muscu
255	21.2	50.5	10977	10	AB196925	AB196925 Japanese	328	21	50.0	2102	6	BC039217	BC039217 Mus muscu
256	21.2	50.5	10977	10	AB196926	AB196926 Japanese	329	21	50.0	2145	2	CQ413719	CQ413719 Sequence
257	21.2	50.5	10977	10	AF069076	AF069076 Japanese	330	21	50.0	2174	6	BC036173	BC036173 Mus muscu
258	21.2	50.5	10977	10	AF080251	AF080251 Japanese	331	21	50.0	2420	6	AY079517S2	AY079517 Mus muscu
259	21.2	50.5	10977	10	AF486638	AF486638 Japanese	C 332	21	50.0	12027	15	AE008569	AE008569 Rickettsi
260	21.2	50.5	30221	12	AC014454	AC014454 Drosophi	C 333	21	50.0	96589	2	AX695599	AX695599 Sequence
261	21.2	50.5	30508	13	TRBSAG	L20156 Trypanosoma	334	21	50.0	103667	12	AC178833	AC178833 Strongylo
262	21.2	50.5	40592	5	HSL191F1	Z68756 Human DNA s	335	21	50.0	110000	15	CP000076_14	Continuation (15 o
263	21.2	50.5	56016	12	AC091679	AC091679 Mus muscu	C 336	21	50.0	125517	5	HS191156	AL009050 Human DNA
264	21.2	50.5	84265	11	AB028136	AB028136 Gallus ga	C 337	21	50.0	126080	5	AY029472	AY029472 Homo sapi
265	21.2	50.5	93871	12	AC165216	AC165216 Medicago	C 338	21	50.0	137723	5	AC110814	AC110814 Homo sapi
266	21.2	50.5	123706	13	TH25N7	AL671259 Trypanoso	C 339	21	50.0	151019	12	AC151938	AC151938 Dasyapus n
267	21.2	50.5	137759	13	TBN19R2	AL671256 Trypanoso	340	21	50.0	152680	12	AC173490	AC173490 Bos tauru
268	21.2	50.5	148365	12	AC165219	AC165219 Medicago	341	21	50.0	158793	12	AC169239	AC169239 Bos tauru
269	21.2	50.5	157119	5	AL3390065	AL3390065 Human DNA	342	21	50.0	159791	5	AC009951	AC009951 Homo sapi
270	21.2	50.5	158822	12	AC136174	AC136174 Rattus no	C 343	21	50.0	164281	11	CR318673	CR318673 Zebrafish
271	21.2	50.5	161807	12	BX936424	BX936424 Danio rer	C 344	21	50.0	166469	12	CT009633	CT009633 Danio rer
272	21.2	50.5	163323	6	AC091395	AC091395 Mus muscu	345	21	50.0	169516	12	AC013934	AC013934 Drosophill
273	21.2	50.5	163690	12	AC140908	AC140908 Homo sapi	C 346	21	50.0	170662	13	AC009905	AC009905 Drosophill
274	21.2	50.5	164238	5	AC106745	AC106745 Homo sapi	C 347	21	50.0	175795	12	AC155028	AC155028 Bos tauru
275	21.2	50.5	164271	6	AC122515	AC122515 Mus muscu	348	21	50.0	182128	11	AC147446	AC147446 Gallus ga
276	21.2	50.5	164386	13	AC006491	AC006491 Drosophill	349	21	50.0	184341	12	AC135835	AC135835 Papio anu
277	21.2	50.5	170298	13	AE003690	AE003690 Drosophill	C 350	21	50.0	190196	12	AC129837	AC129837 Papio anu
278	21.2	50.5	173561	12	AC016809	AC016809 Homo sapi	351	21	50.0	194763	12	AC169116	AC169116 Bos tauru
279	21.2	50.5	174471	6	AC122534	AC122534 Mus muscu	C 352	21	50.0	195466	12	AC176508	AC176508 Strongylo
280	21.2	50.5	192226	5	AC109000	AC109000 Rattus no	353	21	50.0	199299	6	AC131718	AC131718 Mus muscu
281	21.2	50.5	192758	5	AC104434	AC104434 Homo sapi	C 354	21	50.0	202399	5	AC090966	AC090966 Papio anu
282	21.2	50.5	202844	6	AC110211	AC110211 Mus muscu	C 355	21	50.0	206846	12	AC165513	AC165513 Bos tauru
283	21.2	50.5	203133	12	AC138168	AC138168 Sus scrof	C 356	21	50.0	213473	6	AC140287	AC140287 Mus muscu
284	21.2	50.5	204001	12	AC148896	AC148896 Otlemur	C 357	21	50.0	216248	12	AC168788	AC168788 Strongylo
285	21.2	50.5	206098	12	AC156915	AC156915 Bos tauru	C 358	21	50.0	217742	12	AC172111	AC172111 Macaca mu
286	21.2	50.5	207639	6	AL844591	AL844591 Mouse DNA	C 359	21	50.0	226088	12	AC152967	AC152967 Bos tauru
287	21.2	50.5	208709	12	AC154058	AC154058 Otlemur	360	21	50.0	234563	12	AC123093	AC123093 Rattus no
288	21.2	50.5	217359	12	AC151059	AC151059 Bos tauru	361	21	50.0	244141	12	AC096072	AC096072 Rattus no
289	21.2	50.5	217591	12	AC130504	AC130504 Rattus no	C 362	21	50.0	258290	12	AC109715	AC109715 Rattus no
290	21.2	50.5	220691	6	AL672270	AL672270 Mouse DNA	C 363	21	50.0	258558	13	AE003654	AE003654 Drosophill
291	21.2	50.5	226284	6	AC099609	AC099609 Mus muscu	364	21	50.0	264683	12	AC105646	AC105646 Rattus no
292	21.2	50.5	231604	12	AC130095	AC130095 Rattus no	365	21	50.0	267565	12	AC098917	AC098917 Rattus no
293	21.2	50.5	244899	12	AC110931	AC110931 Rattus no	366	21	50.0	300953	12	AC162349	AC162349 Bos tauru
294	21.2	50.5	245387	12	AL358052	AL358052 Homo sapi	C 367	21	50.0	334392	12	AC118975	AC118975 Rattus no
295	21.2	50.5	247418	12	AC107347	AC107347 Rattus no	C 368	21	50.0	348764	15	BX569689	BX569689 Synchoco
296	21.2	50.5	253919	12	AC109035	AC109035 Rattus no	C 369	20.8	49.5	917	14	DQ240470	DQ240470 Nyctereut
297	21.2	50.5	286014	12	AC112125	AC112125 Rattus no	370	20.8	49.5	1728	13	MSU45682	AR348475 Sequence
298	21.2	50.5	286014	12	AC112125	AC112125 Bos tauru	371	20.8	49.5	1832	13	AR348475	U46682 Manduca sex
299	21.2	50.5	302156	12	AC157159	AC157159 Bos tauru	372	20.8	49.5	1935	5	AK098749	AK098749 Homo sapi
300	21.2	50.5	306003	12	AC110456	AC110456 Rattus no	373	20.8	49.5	4001	5	AC133795	AC133795 Homo sapi
301	21	50.0	251	15	AY345080	AY345080 Rickettsi	374	20.8	49.5	5565	5	AF166009S3	AF166011 Homo sapi
302	21	50.0	428	11	AY844463	AY844463 Hyla taen	375	20.8	49.5	89295	12	AC159736	AC159736 Homo sapi
303	21	50.0	527	2	CQ396550	CQ396550 Sequence	376	20.8	49.5	101034	5	AC096550	AC096550 Homo sapi
304	21	50.0	527	2	CQ402864	CQ402864 Sequence	C 377	20.8	49.5	110000	12	PFMAL8P1_05	Continuation (6 of
305	21	50.0	558	7	BV586363	BV586363 G591P6319	378	20.8	49.5	110000	15	AP007255_30	Continuation (31 o
306	21	50.0	559	2	CQ398658	CQ398658 Sequence	379	20.8	49.5	110000	15	AP007255_31	Continuation (32 o
307	21	50.0	579	2	CQ404981	CQ404981 Sequence	C 380	20.8	49.5	110828	12	AC087586	AC087586 Homo sapi
308	21	50.0	579	2	CQ397122	CQ397122 Sequence	C 381	20.8	49.5	114950	5	AC097652	AC097652 Homo sapi
309	21	50.0	620	7	BV048212	BV048212 S212P6044	C 382	20.8	49.5	120769	14	CR956354	CR956354 Wallaby D
310	21	50.0	625	2	CQ411365	CQ411365 Sequence	383	20.8	49.5	123716	12	CT027747	Danio rer

C 384	20.8	49.5	126692	5	AL450997	AL450997 Human DNA	457	20.6	49.0	2085	2	CQ490219	CQ490219 Sequence
C 385	20.8	49.5	130609	12	AC091798	AC091798 Felis cat	458	20.6	49.0	2085	2	CQ490316	CQ490316 Sequence
C 386	20.8	49.5	133377	12	CT025853	CT025853 Danio rer	459	20.6	49.0	2085	2	CQ496081	CQ496081 Sequence
C 387	20.8	49.5	139076	5	AC004147	AC004147 Homo sapi	460	20.6	49.0	2085	2	CQ496175	CQ496175 Sequence
C 388	20.8	49.5	142988	11	AC145797	AC145797 Xenopus t	461	20.6	49.0	2407	5	BC000770	BC000770 Homo sapi
C 389	20.8	49.5	144015	12	AL357144	AL357144 Homo sapi	c 462	20.6	49.0	2483	15	AB126566	AB126566 Bacillus
C 390	20.8	49.5	146437	12	AC022836	AC022836 Homo sapi	c 463	20.6	49.0	2610	2	BD244712	BD244712 Genes enc
C 391	20.8	49.5	148875	12	AC0162752	AC0162752 Loxodonta	464	20.6	49.0	2610	2	AX023369	AX023369 Sequence
C 392	20.8	49.5	150214	5	AC011291	AC011291 Homo sapi	465	20.6	49.0	2725	5	BC014489	BC014489 Homo sapi
C 393	20.8	49.5	152350	12	AC107322	AC107322 Felis cat	466	20.6	49.0	2759	2	CS130835	CS130835 Sequence
C 394	20.8	49.5	156082	11	BX936402	BX936402 Zebrafish	467	20.6	49.0	2767	5	BC004237	BC004237 Homo sapi
C 395	20.8	49.5	159440	12	AC079012	AC079012 Homo sapi	468	20.6	49.0	2772	2	BD157094	BD157094 Primer fo
C 396	20.8	49.5	160114	12	AC022181	AC022181 Homo sapi	469	20.6	49.0	2772	2	AX878224	AX878224 Sequence
C 397	20.8	49.5	161269	12	AC169522	AC169522 Macaca mu	470	20.6	49.0	2772	5	AK002127	AK002127 Homo sapi
C 398	20.8	49.5	163352	5	AC009225	AC009225 Mus muscu	c 471	20.6	49.0	2977	2	CS130802	CS130802 Sequence
C 399	20.8	49.5	164204	6	AC134550	AC134550 Homo sapi	c 472	20.6	49.0	4279	5	HS9800933	HS9800933
C 400	20.8	49.5	166348	12	CR847851	CR847851 Danio rer	c 473	20.6	49.0	4403	5	BC052980	BC052980 Homo sapi
C 401	20.8	49.5	166906	11	AC146480	AC146480 Danio rer	c 474	20.6	49.0	7013	5	AB002331	AB002331 Homo sapi
C 402	20.8	49.5	167320	12	AC172719	AC172719 Callicebu	475	20.6	49.0	7548	5	AY481571	AY481571 Homo sapi
C 403	20.8	49.5	170597	11	AL929558	AL929558 Zebrafish	476	20.6	49.0	8466	5	AY481572	AY481572 Homo sapi
C 404	20.8	49.5	171421	14	AC124909	AC124909 Equus cab	c 477	20.6	49.0	19839	5	AF094481	AF094481 Homo sapi
C 405	20.8	49.5	172560	12	AC091794	AC091794 Felis cat	c 478	20.6	49.0	25001	5	AF490407	AF490407 Homo sapi
C 406	20.8	49.5	172671	5	AC114401	AC114401 Homo sapi	479	20.6	49.0	110000	15	CP000152	CP000152 25
C 407	20.8	49.5	175583	12	AC083850	AC083850 Homo sapi	c 480	20.6	49.0	117669	12	AC171771	AC171771 Sorex ara
C 408	20.8	49.5	176949	6	AC159641	AC159641 Mus muscu	c 481	20.6	49.0	119038	12	AC167321	AC167321 Sorex ara
C 409	20.8	49.5	177332	11	AC148466	AC148466 Xenopus t	482	20.6	49.0	122416	5	AL451050	AL451050 Human DNA
C 410	20.8	49.5	178389	5	AL731555	AL731555 Human DNA	483	20.6	49.0	146343	5	AL590131	AL590131 Human DNA
C 411	20.8	49.5	179510	12	AC013820	AC013820 Homo sapi	484	20.6	49.0	146515	5	AC104009	AC104009 Homo sapi
C 412	20.8	49.5	180092	6	CT009486	CT009486 Mouse DNA	c 485	20.6	49.0	147079	5	AL772262	AL772262 Human DNA
C 413	20.8	49.5	180398	12	AC115766	AC115766 Rattus no	c 486	20.6	49.0	147103	12	AC164508	AC164508 Loxodonta
C 414	20.8	49.5	182859	12	AL845341	AL845341 Mus muscu	c 487	20.6	49.0	152867	5	AC027671	AC027671 Homo sapi
C 415	20.8	49.5	183431	12	AC138788	AC138788 Sus scrof	488	20.6	49.0	154803	5	AC119733	AC119733 Homo sapi
C 416	20.8	49.5	186217	6	AC122310	AC122310 Mus muscu	489	20.6	49.0	155182	5	AC074093	AC074093 Homo sapi
C 417	20.8	49.5	187423	13	AC167369	AC167369 Bos tauru	c 490	20.6	49.0	160241	5	HS98517	HS98517 Human DNA
C 418	20.8	49.5	188905	6	AC084323	AC084323 Mus muscu	491	20.6	49.0	163924	12	AL162592	AL162592 Homo sapi
C 419	20.8	49.5	190191	12	AC138784	AC138784 Sus scrof	492	20.6	49.0	165768	12	AL672011	AL672011 Mus muscu
C 420	20.8	49.5	192010	6	AL732320	AL732320 Mouse DNA	c 493	20.6	49.0	168025	12	AC036200	AC036200 Homo sapi
C 421	20.8	49.5	194536	6	AL833786	AL833786 Mouse DNA	c 494	20.6	49.0	168662	5	CNS01BIS	CNS01BIS Human chr
C 422	20.8	49.5	195229	12	AC155968	AC155968 Xenopus t	c 495	20.6	49.0	167336	12	AC0211350	AC0211350 Homo sapi
C 423	20.8	49.5	201685	12	AC165978	AC165978 Macaca mu	c 496	20.6	49.0	168076	12	AC013304	AC013304 Homo sapi
C 424	20.8	49.5	204952	12	AC148618	AC148618 Salmo sal	c 497	20.6	49.0	168243	12	AC011269	AC011269 Homo sapi
C 425	20.8	49.5	209355	12	AC123163	AC123163 Rattus no	c 498	20.6	49.0	168370	6	AC147104	AC147104 Mus muscu
C 426	20.8	49.5	210374	12	AC010199	AC010199 Homo sapi	c 499	20.6	49.0	168462	11	BX005308	BX005308 Zebrafish
C 427	20.8	49.5	215972	5	AC119543	AC119543 Rattus no	500	20.6	49.0	172052	6	AL645747	AL645747 Mouse DNA
C 428	20.8	49.5	217346	5	AC012363	AC012363 Homo sapi	501	20.6	49.0	172650	12	AC011042	AC011042 Homo sapi
C 429	20.8	49.5	222138	12	AC093961	AC093961 Rattus no	502	20.6	49.0	172779	5	AC093700	AC093700 Homo sapi
C 430	20.8	49.5	222679	12	AC124844	AC124844 Rattus no	c 503	20.6	49.0	173119	6	AC122460	AC122460 Mus muscu
C 431	20.8	49.5	223839	12	CR769780	CR769780 Danio rer	504	20.6	49.0	179104	12	AC172076	AC172076 Bos tauru
C 432	20.8	49.5	227857	12	AC110685	AC110685 Rattus no	c 505	20.6	49.0	179388	12	AC150060	AC150060 Gallus ga
C 433	20.8	49.5	230301	12	AC115521	AC115521 Rattus no	c 506	20.6	49.0	185197	12	AC159026	AC159026 Pan trogl
C 434	20.8	49.5	232679	12	AC130866	AC130866 Rattus no	c 507	20.6	49.0	190773	10	AY261365	AY261365 African s
C 435	20.8	49.5	235225	12	AC113810	AC113810 Rattus no	508	20.6	49.0	200822	5	AL591806	AL591806 Human DNA
C 436	20.8	49.5	235424	12	AC130866	AC130866 Rattus no	509	20.6	49.0	203548	12	AC102094	AC102094 Mus muscu
C 437	20.8	49.5	244470	12	BX571722	BX571722 Danio rer	c 510	20.6	49.0	211451	12	AC009883	AC009883 Homo sapi
C 438	20.8	49.5	246101	12	AC182068	AC182068 Bos tauru	c 511	20.6	49.0	218619	12	AC108645	AC108645 Rattus no
C 439	20.8	49.5	247167	12	AC132985	AC132985 Rattus no	c 512	20.6	49.0	218892	5	AC007225	AC007225 Homo sapi
C 440	20.8	49.5	252650	12	AC160208	AC160208 Bos tauru	c 513	20.6	49.0	220242	6	AL603706	AL603706 Mouse DNA
C 441	20.8	49.5	255305	12	AC160576	AC160576 Bos tauru	514	20.6	49.0	222028	12	AC097167	AC097167 Rattus no
C 442	20.8	49.5	256482	12	AC131342	AC131342 Mus muscu	c 515	20.6	49.0	227876	12	AC097192	AC097192 Rattus no
C 443	20.8	49.5	258542	12	AC129865	AC129865 Rattus no	516	20.6	49.0	230079	12	AC159216	AC159216 Pan trogl
C 444	20.8	49.5	268208	12	AC106249	AC106249 Rattus no	c 517	20.6	49.0	236629	12	AC120586	AC120586 Rattus no
C 445	20.8	49.5	268524	12	AC106223	AC106223 Rattus no	c 518	20.6	49.0	241684	12	AC125764	AC125764 Rattus no
C 446	20.8	49.5	292409	11	BX530407	BX530407 Zebrafish	519	20.6	49.0	245743	12	AC125999	AC125999 Rattus no
C 447	20.6	49.0	170	2	CQ706799	CQ706799 Sequence	c 520	20.6	49.0	246832	12	AC094056	AC094056 Rattus no
C 448	20.6	49.0	330	7	G62184	G62184 EST543 Foet	c 521	20.6	49.0	257334	12	AC096266	AC096266 Rattus no
C 449	20.6	49.0	428	11	AY844405	AY844405 Hyla cavi	522	20.6	49.0	259141	12	AC117864	AC117864 Rattus no
C 450	20.6	49.0	617	5	IP0EST087	IP0EST087 Homo sapi	c 523	20.6	49.0	260427	12	AC152740	AC152740 Bos tauru
C 451	20.6	49.0	670	7	BV483698	BV483698 S21P6835	c 524	20.6	49.0	275125	12	AC114842	AC114842 Rattus no
C 452	20.6	49.0	682	2	CQ412752	CQ412752 Sequence	525	20.6	49.0	286145	12	AC099222	AC099222 Rattus no
C 453	20.6	49.0	701	7	BV474563	BV474563 G591P6270	c 526	20.6	49.0	312348	12	AC160676	AC160676 Bos tauru
C 454	20.6	49.0	901	4	BT017508	BT017508 Zea mays	527	20.6	49.0	325387	12	AC107616	AC107616 Homo sapi
C 455	20.6	49.0	1309	11	BC075213	BC075213 Xenopus 1	c 528	20.4	48.6	264	2	AR273081	AR273081 Sequence
C 456	20.6	49.0	1700	4	BT018724	BT018724 Zea mays	c 529	20.4	48.6	264	2	AR276662	AR276662 Sequence

C 530	20.4	48.6	264	2	AR406937	Sequence	AR406937	C 603	20.4	48.6	162861	12	AC018597	AC018597 Homo sapi
C 531	20.4	48.6	264	2	AR440787	Sequence	AR440787	C 604	20.4	48.6	164168	5	AL135927	AL135927 Human DNA
C 532	20.4	48.6	264	2	AR472945	Sequence	AR472945	C 605	20.4	48.6	164179	5	AC007227	AC007227 Homo sapi
C 533	20.4	48.6	264	2	AR543598	Sequence	AR543598	C 606	20.4	48.6	164450	13	AC011763	AC011763 Drosophila
C 534	20.4	48.6	264	2	AR638796	Sequence	AR638796	C 607	20.4	48.6	164955	12	AC131864	AC131864 Rattus no
C 535	20.4	48.6	264	2	AR695429	Sequence	AR695429	C 608	20.4	48.6	167647	6	AC144913	AC144913 Mus muscu
C 536	20.4	48.6	264	2	AX063196	Sequence	AX063196	C 609	20.4	48.6	168528	6	AC132255	AC132255 Mus muscu
C 537	20.4	48.6	264	2	AX368113	Sequence	AX368113	C 610	20.4	48.6	169085	6	AC102666	AC102666 Mus muscu
C 538	20.4	48.6	363	2	AR622535	Sequence	AR622535	C 611	20.4	48.6	171875	12	AC099746	AC099746 Bos tauru
C 539	20.4	48.6	540	7	AU027881	Rattus no	AU027881	C 612	20.4	48.6	172382	12	AC149046	AC149046 Papio anu
C 540	20.4	48.6	615	7	BV352364	Sequence	BV352364	C 613	20.4	48.6	172659	12	AC024320	AC024320 Homo sapi
C 541	20.4	48.6	629	7	CQ523848	Sequence	CQ523848	C 614	20.4	48.6	172755	5	CNS01RGL	AL158111 Human chr
C 542	20.4	48.6	703	7	BV052360	Sequence	BV052360	C 615	20.4	48.6	173403	12	AC164861	AC164861 Bos tauru
C 543	20.4	48.6	707	2	AX436879	Sequence	AX436879	C 616	20.4	48.6	173891	12	BX927251	BX927251 Danio rer
C 544	20.4	48.6	708	2	AX436719	Sequence	AX436719	C 617	20.4	48.6	176105	12	AC153235	AC153235 Bos tauru
C 545	20.4	48.6	728	2	BD148430	Primer fo	BD148430	C 618	20.4	48.6	176399	5	CNS05TE1	AL359398 Human chr
C 546	20.4	48.6	728	2	AX668368	Sequence	AX668368	C 619	20.4	48.6	177041	6	AC126024	AC126024 Mus muscu
C 547	20.4	48.6	770	11	CR405907	Sequence	CR405907	C 620	20.4	48.6	177273	5	AC074194	AC074194 Homo sapi
C 548	20.4	48.6	1356	2	CQ717415	Sequence	CQ717415	C 621	20.4	48.6	179576	12	AL170664	AL170664 Bos tauru
C 549	20.4	48.6	1752	14	BC110126	Bos tauru	BC110126	C 622	20.4	48.6	179738	5	AL354938	AL354938 Human DNA
C 550	20.4	48.6	2523	2	AX714593	Sequence	AX714593	C 623	20.4	48.6	181120	6	AC161148	AC161148 Mus muscu
C 551	20.4	48.6	2523	5	AK057043	Homo sapi	AK057043	C 624	20.4	48.6	181782	6	AC161172	AC161172 Mus muscu
C 552	20.4	48.6	2648	2	AR717849	Sequence	AR717849	C 625	20.4	48.6	188137	12	AC148796	AC148796 Otollemur
C 553	20.4	48.6	2648	2	AX746824	Sequence	AX746824	C 626	20.4	48.6	193044	12	AC027769	AC027769 Homo sapi
C 554	20.4	48.6	2648	5	AK091168	Homo sapi	AK091168	C 627	20.4	48.6	193312	12	AC157839	AC157839 Otollemur
C 555	20.4	48.6	2655	5	BC103826	Homo sapi	BC103826	C 628	20.4	48.6	194339	6	AC109199	AC109199 Mus muscu
C 556	20.4	48.6	2655	5	BC103827	Homo sapi	BC103827	C 629	20.4	48.6	196972	12	AC166531	AC166531 Bos tauru
C 557	20.4	48.6	2737	2	CQ606770	Sequence	CQ606770	C 630	20.4	48.6	197273	6	AC024607	AC024607 Mus muscu
C 558	20.4	48.6	2952	5	BC008928	Homo sapi	BC008928	C 631	20.4	48.6	198777	11	AL807749	AL807749 Zebrafish
C 559	20.4	48.6	2970	5	BC064624	Homo sapi	BC064624	C 632	20.4	48.6	205550	12	AL148945	AL148945 Otollemur
C 560	20.4	48.6	2973	5	BC075854	Homo sapi	BC075854	C 633	20.4	48.6	208470	6	AL928883	AL928883 Mouse DNA
C 561	20.4	48.6	2998	5	HS4409349	Homo sapi	HS4409349	C 634	20.4	48.6	214809	12	AC099407	AC099407 Bos tauru
C 562	20.4	48.6	3057	2	CS245209	Sequence	CS245209	C 635	20.4	48.6	216158	12	AC113563	AC113563 Rattus no
C 563	20.4	48.6	3240	2	BD156445	Primer fo	BD156445	C 636	20.4	48.6	216289	12	AC123712	AC123712 Mus muscu
C 564	20.4	48.6	3240	2	AX877025	Sequence	AX877025	C 637	20.4	48.6	216615	6	AC107850	AC107850 Mus muscu
C 565	20.4	48.6	3240	5	AK001574	Homo sapi	AK001574	C 638	20.4	48.6	217725	13	AE003717	AE003717 Drosophila
C 566	20.4	48.6	3722	5	AK124755	Homo sapi	AK124755	C 639	20.4	48.6	218524	6	AC151721	AC151721 Mus muscu
C 567	20.4	48.6	5336	11	AB097142	Danio rer	AB097142	C 640	20.4	48.6	224009	12	AC164301	AC164301 Mus muscu
C 568	20.4	48.6	6172	15	MTMCR	Sequence	MTMCR	C 641	20.4	48.6	224101	6	AC145301	AC145301 Mus muscu
C 569	20.4	48.6	21295	2	AR619799	Sequence	AR619799	C 642	20.4	48.6	227407	12	AC170467	AC170467 Bos tauru
C 570	20.4	48.6	56179	4	AP004626	Lotus cor	AP004626	C 643	20.4	48.6	228155	6	AC164411	AC164411 Mus muscu
C 571	20.4	48.6	67395	12	AC134417	Mus muscu	AC134417	C 644	20.4	48.6	229761	12	AC123314	AC123314 Rattus no
C 572	20.4	48.6	67971	5	AL138124	Homo sapi	AL138124	C 645	20.4	48.6	230187	12	AC152198	AC152198 Bos tauru
C 573	20.4	48.6	85763	5	AL359455	Human DNA	AL359455	C 646	20.4	48.6	230212	6	CT025668	CT025668 Mouse DNA
C 574	20.4	48.6	92969	6	BX842667	Mouse DNA	BX842667	C 647	20.4	48.6	233137	12	CR293502	CR293502 Danio rer
C 575	20.4	48.6	94135	12	AC168633	Strongylo	AC168633	C 648	20.4	48.6	241100	12	AC095453	AC095453 Rattus no
C 576	20.4	48.6	110000	2	BD430793_16	Continuation (17 o	BD430793_16	C 649	20.4	48.6	241958	12	AC162594	AC162594 Bos tauru
C 577	20.4	48.6	110000	4	AP008213_236	Continuation (237	AP008213_236	C 650	20.4	48.6	247017	12	AC110624	AC110624 Rattus no
C 578	20.4	48.6	110000	4	AP008214_090	Continuation (91 o	AP008214_090	C 651	20.4	48.6	251796	12	AC094921	AC094921 Rattus no
C 579	20.4	48.6	110000	4	CR382130_25	Continuation (26 o	CR382130_25	C 652	20.4	48.6	251806	12	AC096305	AC096305 Rattus no
C 580	20.4	48.6	110000	4	CR382131_18	Continuation (19 o	CR382131_18	C 653	20.4	48.6	252995	12	AC091638	AC091638 Rattus no
C 581	20.4	48.6	110000	12	AC013623_1	Continuation (2 of	AC013623_1	C 654	20.4	48.6	255019	12	AC162413	AC162413 Bos tauru
C 582	20.4	48.6	110000	12	AC021632_0	Continuation (30 o	AC021632_0	C 655	20.4	48.6	263704	12	AC162592	AC162592 Bos tauru
C 583	20.4	48.6	110000	15	CR954246_29	Continuation (17 o	CR954246_29	C 656	20.4	48.6	263776	12	AC087335	AC087335 Mus muscu
C 584	20.4	48.6	119235	15	AL011817	Homo sapi	AL011817	C 657	20.4	48.6	264661	12	AC094217	AC094217 Rattus no
C 585	20.4	48.6	122864	5	AL390863	Human DNA	AL390863	C 658	20.4	48.6	265118	2	AX041922	AX041922 Sequence
C 586	20.4	48.6	123787	4	AP003747	Oryza sat	AP003747	C 659	20.4	48.6	265118	15	CNSPAX06	AX248288 Pyrococcus
C 587	20.4	48.6	128318	5	AC097450	Homo sapi	AC097450	C 660	20.4	48.6	276285	12	AC128520	AC128520 Rattus no
C 588	20.4	48.6	129108	12	AC162742	Loxodonta	AC162742	C 661	20.4	48.6	280325	12	AC112293	AC112293 Rattus no
C 589	20.4	48.6	133248	12	AC163974	Loxodonta	AC163974	C 662	20.4	48.6	319246	12	BX546472	BX546472 Danio rer
C 590	20.4	48.6	141473	11	BX942819	Zebrafish	BX942819	C 663	20.2	48.1	384	2	AX537754	AX537754 Sequence
C 591	20.4	48.6	145888	12	AC073219	Homo sapi	AC073219	C 664	20.2	48.1	384	4	MRA400724	AJ400724 Mucor rac
C 592	20.4	48.6	153106	12	AC162469	Bos tauru	AC162469	C 665	20.2	48.1	428	11	AY844365	AY844365 Aplastodi
C 593	20.4	48.6	155544	4	AP005799	Oryza sat	AP005799	C 666	20.2	48.1	428	11	AY844366	AY844366 Aplastodi
C 594	20.4	48.6	156318	5	AC132834	Homo sapi	AC132834	C 667	20.2	48.1	428	11	AY844377	AY844377 Duellmano
C 595	20.4	48.6	157020	12	AC074118	Homo sapi	AC074118	C 668	20.2	48.1	428	11	AY844384	AY844384 Hyla albo
C 596	20.4	48.6	157660	12	AC027774	Homo sapi	AC027774	C 669	20.2	48.1	428	11	AY844385	AY844385 Hyla albo
C 597	20.4	48.6	157905	4	AP005495	Oryza sat	AP005495	C 670	20.2	48.1	428	11	AY844386	AY844386 Hyla ance
C 598	20.4	48.6	159144	12	AC116649	Homo sapi	AC116649	C 671	20.2	48.1	428	11	AY844388	AY844388 Hyla anne
C 599	20.4	48.6	160355	12	AC022644	Homo sapi	AC022644	C 672	20.2	48.1	428	11	AY844392	AY844392 Hyla aril
C 600	20.4	48.6	160561	12	AC017724	Drosophila	AC017724	C 673	20.2	48.1	428	11	AY844394	AY844394 Hyla aviv
C 601	20.4	48.6	160990	5	AL157702	Human DNA	AL157702	C 674	20.2	48.1	428	11	AY844396	AY844396 Hyla beni
C 602	20.4	48.6	162028	12	AC149108	Papio anu	AC149108	C 675	20.2	48.1	428	11	AY844400	AY844400 Hyla brev

676	20.2	48.1	428	11	AY844402	AY844402 Hyia call	749	20.2	48.1	110000	15	AE000657_00	AE000657 Aquifex a
677	20.2	48.1	428	11	AY844404	AY844404 Hyia carn	c 750	20.2	48.1	110000	15	CP000103_06	Continuation (7 of
678	20.2	48.1	428	11	AY844408	AY844408 Hyia cine	751	20.2	48.1	110000	15	CP000112_06	Continuation (7 of
679	20.2	48.1	428	11	AY844409	AY844409 Hyia circ	752	20.2	48.1	110000	15	AP0008384_04	Continuation (5 of
680	20.2	48.1	428	11	AY844410	AY844410 Hyia coly	c 753	20.2	48.1	111868	15	AC180846	AC180846 Strongylo
681	20.2	48.1	428	11	AY844417	AY844417 Hyia gies	754	20.2	48.1	113915	5	AC092078	AC092078 Homo sapi
682	20.2	48.1	428	11	AY844418	AY844418 Hyia grat	c 755	20.2	48.1	116194	12	AC172875	AC172875 Brassica
683	20.2	48.1	428	11	AY844419	AY844419 Hyia hyla	756	20.2	48.1	119420	12	AC027485	AC027485 Homo sapi
684	20.2	48.1	428	11	AY844423	AY844423 Hyia lema	c 757	20.2	48.1	123009	12	AC181492	AC181492 Strongylo
685	20.2	48.1	428	11	AY844425	AY844425 Hyia leuc	c 758	20.2	48.1	123291	5	AC104070	AC104070 Homo sapi
686	20.2	48.1	428	11	AY844430	AY844430 Hyia micr	759	20.2	48.1	124552	5	AC0011403	AC0011403 Homo sapi
687	20.2	48.1	428	11	AY844435	AY844435 Hyia miya	c 760	20.2	48.1	131152	6	AL929068	AL929068 Mouse DNA
688	20.2	48.1	428	11	AY844436	AY844436 Hyia mult	c 761	20.2	48.1	134420	5	AC010428	AC010428 Homo sapi
689	20.2	48.1	428	11	AY844439	AY844439 Hyia palm	c 762	20.2	48.1	134940	12	AC018939	AC018939 Homo sapi
690	20.2	48.1	428	11	AY844440	AY844440 Hyia parv	763	20.2	48.1	135028	12	AC181466	AC181466 Strongylo
691	20.2	48.1	428	11	AY844444	AY844444 Hyia pseu	764	20.2	48.1	136853	6	AC165307	AC165307 Mus muscu
692	20.2	48.1	428	11	AY844446	AY844446 Hyia rhod	c 765	20.2	48.1	138159	5	AL138965	AL138965 Human DNA
693	20.2	48.1	428	11	AY844451	AY844451 Hyia sara	c 766	20.2	48.1	140042	6	AC157909	AC157909 Mus muscu
694	20.2	48.1	428	11	AY844455	AY844455 Hyia sibi	c 767	20.2	48.1	143092	13	CEY1797B	AL023828 Caenorhab
695	20.2	48.1	428	11	AY844456	AY844456 Hyia aff.	c 768	20.2	48.1	143461	12	AC141021	AC141021 Rattus no
696	20.2	48.1	428	11	AY844457	AY844457 Hyia sp.	769	20.2	48.1	146699	5	BS000226	BS000226 Pan trogl
697	20.2	48.1	428	11	AY844458	AY844458 Hyia sp.	770	20.2	48.1	146891	12	AL845302	AL845302 Oryctolag
698	20.2	48.1	428	11	AY844461	AY844461 Hyia aff.	c 771	20.2	48.1	147219	11	AL845302	AL845302 Zebrafish
699	20.2	48.1	428	11	AY844467	AY844467 Hyia weyg	772	20.2	48.1	148628	12	AC157451	AC157451 Dasyypus n
700	20.2	48.1	428	11	AY844478	AY844478 Melanophr	773	20.2	48.1	148892	4	AP003235	AP003235 Oryza sat
701	20.2	48.1	428	11	AY844481	AY844481 Osteoceph	774	20.2	48.1	152523	5	AC007009	AC007009 Homo sapi
702	20.2	48.1	428	11	AY844483	AY844483 Osteoceph	c 775	20.2	48.1	154524	6	AC132278	AC132278 Mus muscu
703	20.2	48.1	428	11	AY844484	AY844484 Osteoceph	776	20.2	48.1	156381	12	AC171966	AC171966 Bos tauru
704	20.2	48.1	428	11	AY844485	AY844485 Osteoceph	c 777	20.2	48.1	157295	12	AC079055	AC079055 Homo sapi
705	20.2	48.1	428	11	AY844490	AY844490 Phrynohya	c 778	20.2	48.1	159469	6	AC121920	AC121920 Mus muscu
706	20.2	48.1	428	11	AY844493	AY844493 Phrynohya	c 779	20.2	48.1	162020	12	AC009864	AC009864 Homo sapi
707	20.2	48.1	428	11	AY844518	AY844518 Scinax el	c 780	20.2	48.1	163755	6	AC166570	AC166570 Mus muscu
708	20.2	48.1	428	11	AY844520	AY844520 Scinax el	c 781	20.2	48.1	164455	6	AC102466	AC102466 Mus muscu
709	20.2	48.1	428	11	AY844521	AY844521 Scinax ru	782	20.2	48.1	164686	5	CNS057CG	AL350597 Human chr
710	20.2	48.1	428	11	AY844522	AY844522 Scinax sq	783	20.2	48.1	165030	12	AC179774	AC179774 Strongylo
711	20.2	48.1	428	11	AY844523	AY844523 Scinax st	c 784	20.2	48.1	165030	12	AC179774	AC179774 Strongylo
712	20.2	48.1	428	11	AY844530	AY844530 Tepuihya	c 785	20.2	48.1	165175	12	AC168383	AC168383 Strongylo
713	20.2	48.1	428	11	AY844531	AY844531 Trachycep	c 786	20.2	48.1	165880	5	AL136372	AL136372 Human DNA
714	20.2	48.1	922	5	AF069967	AF069967 Hylobates	c 787	20.2	48.1	166407	12	AC176375	AC176375 Strongylo
715	20.2	48.1	1018	5	AF0612134	AY012134 Hylobates	788	20.2	48.1	166451	5	AC018994	AC018994 Homo sapi
716	20.2	48.1	1455	4	BT000206	BT000206 Aradidops	789	20.2	48.1	167144	6	AC158748	AC158748 Mus muscu
717	20.2	48.1	1534	11	AY323766	AY323766 Hyia cine	c 790	20.2	48.1	167764	12	AC073870	AC073870 Homo sapi
718	20.2	48.1	1545	2	AX300259	AX300259 Sequence	791	20.2	48.1	170393	12	AC024304	AC024304 Homo sapi
719	20.2	48.1	1557	4	AY057647	AY057647 Aradidops	c 792	20.2	48.1	170936	5	AC147028	AC147028 Pan trogl
720	20.2	48.1	1733	4	AY136392	AY136392 Aradidops	793	20.2	48.1	171790	12	AC011106	AC011106 Homo sapi
721	20.2	48.1	4327	6	BC079620	BC079620 Mus muscu	c 794	20.2	48.1	172576	5	CNS01DSG	AL121775 Human chr
722	20.2	48.1	5919	6	BC095958	BC095958 Mus muscu	795	20.2	48.1	172949	12	AC123431	AC123431 Rattus no
723	20.2	48.1	6005	6	AY043326	AY043326 Rattus no	c 796	20.2	48.1	173094	12	AC141547	AC141547 Rattus no
724	20.2	48.1	6195	6	BC058681	BC058681 Mus muscu	c 797	20.2	48.1	174130	11	AL732628	AL732628 Zebrafish
725	20.2	48.1	6268	6	AK122284	AK122284 Mus muscu	798	20.2	48.1	176190	12	AC180851	AC180851 Strongylo
726	20.2	48.1	6400	6	AF026504	AF026504 Rattus no	799	20.2	48.1	176269	12	AC012663	AC012663 Homo sapi
727	20.2	48.1	6429	6	AK129343	AK129343 Mus muscu	c 800	20.2	48.1	176933	5	AC011994	AC011994 Homo sapi
728	20.2	48.1	16472	5	HLMITCSEQ	X99256 Hylobates l	801	20.2	48.1	177255	12	AC123235	AC123235 Rattus no
729	20.2	48.1	4096	13	U80437	U80437 Caenorhabdi	c 802	20.2	48.1	178830	4	AP003286	AP003286 Oryza sat
730	20.2	48.1	43799	12	AC100909	AC100909 Mus muscu	c 803	20.2	48.1	179491	12	AC020673	AC020673 Homo sapi
731	20.2	48.1	53394	5	AC013276	AC013276 Homo sapi	c 804	20.2	48.1	180363	6	AC154755	AC154755 Mus muscu
732	20.2	48.1	68316	12	CEH16D04	Z92795 Caenorhabdi	c 805	20.2	48.1	181208	6	AC124370	AC124370 Mus muscu
733	20.2	48.1	71260	11	AC151318	AC151318 Xenopus t	c 806	20.2	48.1	182108	12	AC170644	AC170644 Bos tauru
734	20.2	48.1	71612	12	AC173273	AC173273 Bos tauru	c 807	20.2	48.1	183494	12	AC079989	AC079989 Rattus no
735	20.2	48.1	71739	5	AL391360	AL391360 Human DNA	c 808	20.2	48.1	184327	6	AC120859	AC120859 Mus muscu
736	20.2	48.1	77101	5	AC090684	AC090684 Homo sapi	c 809	20.2	48.1	184377	6	AC121579	AC121579 Mus muscu
737	20.2	48.1	79186	12	AC006781	AC006781 Caenorhab	810	20.2	48.1	184439	12	AP003085	AP003085 Homo sapi
738	20.2	48.1	79186	12	AC006859	AC006859 Caenorhab	811	20.2	48.1	184623	6	AC158570	AC158570 Mus muscu
739	20.2	48.1	79452	12	AC016536	AC016536 Homo sapi	812	20.2	48.1	185561	12	AC117124	AC117124 Rattus no
740	20.2	48.1	95449	5	AP001620	AP001620 Homo sapi	c 813	20.2	48.1	185809	12	AC141579	AC141579 Rattus no
741	20.2	48.1	97990	5	AC025167	AC025167 Homo sapi	c 814	20.2	48.1	187451	12	AC176132	AC176132 Strongylo
742	20.2	48.1	103244	14	CR956378	CR956378 Pig DNA s	c 815	20.2	48.1	187568	5	AP002967	AP002967 Homo sapi
743	20.2	48.1	103917	12	AC125501	AC125501 Takifugu	c 816	20.2	48.1	187752	6	AC124689	AC124689 Mus muscu
744	20.2	48.1	105104	12	AC110828	Continuation (7 of	c 817	20.2	48.1	188376	6	AC112267	AC112267 Mus muscu
745	20.2	48.1	105259	5	AC116331	Continuation (7 of	818	20.2	48.1	189137	6	AC079378	AC079378 Rattus no
746	20.2	48.1	105543	4	ATAC009325	AC009325 Aradidops	c 819	20.2	48.1	189847	12	AC091964	AC091964 Homo sapi
747	20.2	48.1	108556	12	AC170551	AC170551 Bos tauru	c 820	20.2	48.1	190374	12	AC132647	AC132647 Rattus no
748	20.2	48.1	110000	4	AP008207_376	Continuation (377	c 821	20.2	48.1	191098	12	AC178716	AC178716 Strongylo

822	20.2	48.1	191747	12	AC113675	AC113675 Rattus no	895	20	47.6	542	2	AX105689	AX105689 Sequence
823	20.2	48.1	193412	12	AC170376	AC170376 Bos tauru	896	20	47.6	546	2	AX319688	AX319688 Sequence
824	20.2	48.1	195269	6	AC153359	AC153359 Mus muscu	897	20	47.6	576	2	AR035907	AR035907 Sequence
825	20.2	48.1	196898	12	AC133945	AC133945 Rattus no	898	20	47.6	576	2	I20143	I20143 Sequence 49
-826	20.2	48.1	198328	12	AC167488	AC167488 Bos tauru	899	20	47.6	576	2	AR340321	AR340321 Sequence
827	20.2	48.1	198392	6	AC022236	AC022236 Mus muscu	900	20	47.6	576	10	HPCPE1R	L16645 Hepatitis C
828	20.2	48.1	199746	12	AC173610	AC173610 Bos tauru	901	20	47.6	624	14	SSSP8C23	L16645 Sus scrof
829	20.2	48.1	199853	12	AC073785	AC073785 Mus muscu	902	20	47.6	649	2	CQ409812	CQ409812 Sequence
830	20.2	48.1	199860	6	AC102131	AC102131 Mus muscu	903	20	47.6	808	7	BV070218	BV070218 S21296040
831	20.2	48.1	200353	6	AC166575	AC166575 Mus muscu	904	20	47.6	1385	4	AK120866	AK120866 Oryza sat
832	20.2	48.1	201694	6	AC111115	AC111115 Mus muscu	905	20	47.6	4636	4	AK120507	AK120507 Oryza sat
833	20.2	48.1	201721	12	AC157290	AC157290 Bos tauru	906	20	47.6	7690	2	AX346025	AX346025 Sequence
834	20.2	48.1	203187	12	CT485788	CT485788 Mus muscu	907	20	47.6	9388	5	AB095937	AB095937 Homo sapi
835	20.2	48.1	205940	12	AC021830	AC021830 Homo sapi	908	20	47.6	24387	5	HSJ185D5	AL118498 Human DNA
836	20.2	48.1	206075	12	AC118511	AC118511 Rattus no	909	20	47.6	38345	5	AY055383	AY055383 Homo sapi
837	20.2	48.1	206383	6	AC008160	AC008160 Mus muscu	910	20	47.6	38703	5	AC000031	AC000031 Homo sapi
838	20.2	48.1	207922	6	AL450399	AL450399 Mouse DNA	911	20	47.6	42659	12	AC165608	AC165608 Bos tauru
839	20.2	48.1	208332	6	AL450399	AL450399 Mouse DNA	912	20	47.6	57358	15	AY803022	AY803022 Neisseria
840	20.2	48.1	209432	12	AC111308	AC111308 Rattus no	913	20	47.6	77227	12	AC025099	AC025099 Homo sapi
841	20.2	48.1	209885	12	AC172289	AC172289 Bos tauru	914	20	47.6	78794	4	OSJN00123	AL606996 Oryza sat
842	20.2	48.1	209999	12	AC019128	AC019128 Homo sapi	915	20	47.6	82767	11	AL672065	AL672065 Zebrafish
843	20.2	48.1	212104	5	AC012497	AC012497 Homo sapi	916	20	47.6	97572	4	AP004155	AP004155 Oryza sat
844	20.2	48.1	213047	12	AC020557	AC020557 Homo sapi	917	20	47.6	99766	12	AL353702	AL353702 Homo sapi
845	20.2	48.1	213181	12	AC122949	AC122949 Rattus no	918	20	47.6	100999	10	AF270937	AF270937 Plutella
846	20.2	48.1	213656	6	AC166172	AC166172 Mus muscu	919	20	47.6	102606	12	AC176436	AC176436 Strongylo
847	20.2	48.1	213669	6	AC144940	AC144940 Mus muscu	920	20	47.6	104353	5	HS1170K4	AL022314 Human DNA
848	20.2	48.1	214372	12	AC148959	AC148959 Otolomur	921	20	47.6	108514	12	AC084722	AC084722 Homo sapi
849	20.2	48.1	216380	6	AC113113	AC113113 Mus muscu	922	20	47.6	109540	4	AC147364	AC147364 Medicago
850	20.2	48.1	219599	12	AC109416	AC109416 Rattus no	923	20	47.6	110000	4	AP008214	Continuation 195 o
851	20.2	48.1	222545	12	AC109051	AC109051 Rattus no	924	20	47.6	110000	4	AP008210	Continuation 185 o
852	20.2	48.1	226153	6	AC074042	AC074042 Mus muscu	925	20	47.6	110000	15	BA000040	Continuation (20 o
853	20.2	48.1	226153	12	AC073704	AC073704 Mus muscu	926	20	47.6	117580	12	AP006455	AP006455 Oryza sat
854	20.2	48.1	227202	12	AC128501	AC128501 Rattus no	927	20	47.6	121581	12	AC167325	AC167325 Loxodonta
855	20.2	48.1	227492	6	AC116585	AC116585 Mus muscu	928	20	47.6	123375	4	OSJN00195	AL662993 Oryza sat
856	20.2	48.1	231435	12	AC113813	AC113813 Rattus no	929	20	47.6	129489	6	AL713860	AL713860 Mouse DNA
857	20.2	48.1	231485	12	AC098427	AC098427 Rattus no	930	20	47.6	136224	2	AC139062	AC139062 Mus muscu
858	20.2	48.1	232159	12	AC182457	AC182457 Gallus ga	931	20	47.6	137472	12	AC180955	AC180955 Strongylo
859	20.2	48.1	232483	12	AC129369	AC129369 Rattus no	932	20	47.6	140765	5	AC021301	AC021301 Homo sapi
860	20.2	48.1	233344	12	AC097829	AC097829 Rattus no	933	20	47.6	141628	5	AL391724	AL391724 Human DNA
861	20.2	48.1	234418	12	AC166379	AC166379 Bos tauru	934	20	47.6	143148	6	AL607034	AL607034 Mouse DNA
862	20.2	48.1	234584	12	AC111604	AC111604 Rattus no	935	20	47.6	146976	6	AL711199	AL711199 Mus muscu
863	20.2	48.1	236698	12	AC098997	AC098997 Rattus no	936	20	47.6	151468	6	AC154184	AC154184 Mus muscu
864	20.2	48.1	238953	12	AC166966	AC166966 Mus muscu	937	20	47.6	153938	12	AC168572	AC168572 Strongylo
865	20.2	48.1	239105	12	AC096791	AC096791 Rattus no	938	20	47.6	155595	12	AC078964	AC078964 Homo sapi
866	20.2	48.1	239591	12	AC128432	AC128432 Rattus no	939	20	47.6	158175	5	AC079807	AC079807 Homo sapi
867	20.2	48.1	240250	12	AC154886	AC154886 Bos tauru	940	20	47.6	159311	12	AC141040	AC141040 Rattus no
868	20.2	48.1	243366	12	AC126872	AC126872 Rattus no	941	20	47.6	160023	12	AC154256	AC154256 Mus muscu
869	20.2	48.1	243739	12	AC109679	AC109679 Rattus no	942	20	47.6	162362	6	AC154685	AC154685 Mus muscu
870	20.2	48.1	243770	12	AC123120	AC123120 Rattus no	943	20	47.6	165090	5	AC010655	AC010655 Homo sapi
871	20.2	48.1	245053	12	AC165130	AC165130 Bos tauru	944	20	47.6	166622	11	BX005135	BX005135 Zebrafish
872	20.2	48.1	245339	12	AC121656	AC121656 Rattus no	945	20	47.6	166669	5	AC113191	AC113191 Homo sapi
873	20.2	48.1	248580	12	AC117158	AC117158 Rattus no	946	20	47.6	167691	5	HS406A7	AL023284 Human DNA
874	20.2	48.1	249840	12	AC164639	AC164639 Mus muscu	947	20	47.6	167749	12	AC021489	AC021489 Homo sapi
875	20.2	48.1	251882	12	AC098075	AC098075 Rattus no	948	20	47.6	168863	5	AC011286	AC011286 Homo sapi
876	20.2	48.1	252054	12	AC097119	AC097119 Rattus no	949	20	47.6	170090	6	AC113593	AC113593 Mus muscu
877	20.2	48.1	252825	12	AC132024	AC132024 Rattus no	950	20	47.6	170928	12	AC119121	AC119121 Rattus no
878	20.2	48.1	258276	12	AC130846	AC130846 Rattus no	951	20	47.6	171162	12	AC068684	AC068684 Homo sapi
879	20.2	48.1	280025	12	AC108655	AC108655 Rattus no	952	20	47.6	171663	6	AC151475	AC151475 Mus muscu
880	20.2	48.1	280215	12	AC129648	AC129648 Rattus no	953	20	47.6	172449	5	AC146031	AC146031 Pan trogl
881	20.2	48.1	283806	12	AC167875	AC167875 Bos tauru	954	20	47.6	177724	12	AC073214	AC073214 Homo sapi
882	20.2	48.1	288211	12	AC181676	AC181676 Strongylo	955	20	47.6	177724	12	AC176849	AC176849 Strongylo
883	20.2	48.1	288211	12	AC181676	AC181676 Strongylo	956	20	47.6	180000	15	AF322012	AF322012 Bradyrhiz
884	20.2	48.1	294136	12	AC006901	AC006901 Caenorhab	957	20	47.6	182102	12	AC008048	AC008048 Homo sapi
885	20.2	48.1	309457	12	AC160280	AC160280 Bos tauru	958	20	47.6	182528	5	AC020743	AC020743 Homo sapi
886	20.2	48.1	326336	12	AC046137	AC046137 Homo sapi	959	20	47.6	184292	6	AC123848	AC123848 Mus muscu
887	20.2	48.1	328736	12	AC107298	AC107298 Homo sapi	960	20	47.6	184868	12	AC156769	AC156769 Rhinolph
888	20.2	48.1	336578	5	AP001745	AP001745 Homo sapi	961	20	47.6	185759	12	AC149612	AC149612 Ictalurus
889	20.2	48.1	340525	12	AC117847	AC117847 Rattus no	962	20	47.6	186114	5	AC021621	AC021621 Homo sapi
890	20.2	48.1	345109	12	AC098623	AC098623 Rattus no	963	20	47.6	186534	12	AC142482	AC142482 Rattus no
891	20.2	48.1	349061	15	NNA257220	AL162753 Neisseria	964	20	47.6	190295	6	AC124445	AC124445 Mus muscu
892	20.2	48.1	349648	15	BX957220	BX957220 Methanoco	965	20	47.6	191821	5	AC022616	AC022616 Homo sapi
893	20	47.6	527	2	CQ424285	CQ424285 Sequence	966	20	47.6	192632	6	AC164554	AC164554 Mus muscu
894	20	47.6	531	2	AX319681	AX319681 Sequence	967	20	47.6	192877	6	AC132286	AC132286 Mus muscu

c 968	20	47.6	193670	5	AC026261	AC026261 Homo sapi
969	20	47.6	201361	11	EX927088	EX927088 Zebrafish
c 970	20	47.6	204505	12	AC119344	AC119344 Rattus no
c 971	20	47.6	204836	12	AC119348	AC119348 Rattus no
c 972	20	47.6	205166	12	AC148169	AC148169 Zea mays
c 973	20	47.6	205338	12	AC141855	AC141855 Sus scrofa
974	20	47.6	205946	6	AC160135	AC160135 Mus muscu
975	20	47.6	208408	12	AC179649	AC179649 Strongylo
976	20	47.6	209732	12	AC166400	AC166400 Bos tauru
977	20	47.6	213907	12	CT573134	CT573134 Danio rer
978	20	47.6	215150	12	AC116438	AC116438 Pan trogl
979	20	47.6	218365	12	AC097708	AC097708 Rattus no
980	20	47.6	219086	5	AC010147	AC010147 Homo sapi
c 981	20	47.6	219548	6	AC165153	AC165153 Mus muscu
c 982	20	47.6	223852	6	AC160113	AC160113 Mus muscu
983	20	47.6	226256	6	AC125082	AC125082 Mus muscu
c 984	20	47.6	227057	12	AC182137	AC182137 Bos tauru
c 985	20	47.6	229214	12	CT030153	CT030153 Danio rer
c 986	20	47.6	229438	12	AC107408	AC107408 Rattus no
987	20	47.6	231559	12	AC111306	AC111306 Rattus no
c 988	20	47.6	233670	12	AC126740	AC126740 Rattus no
c 989	20	47.6	238933	12	AC096612	AC096612 Rattus no
c 990	20	47.6	239760	12	AC095113	AC095113 Rattus no
c 991	20	47.6	240425	6	CNS0710T	AL713339 Mus muscu
992	20	47.6	240447	6	CT009630	CT009630 Mouse DNA
993	20	47.6	241178	12	AC130508	AC130508 Rattus no
c 994	20	47.6	242832	12	AC129142	AC129142 Rattus no
c 995	20	47.6	245259	12	AC120898	AC120898 Rattus no
c 996	20	47.6	245808	12	AC114440	AC114440 Rattus no
c 997	20	47.6	246387	12	AC119756	AC119756 Rattus no
998	20	47.6	250175	12	AC160494	AC160494 Bos tauru
c 999	20	47.6	252672	12	AC107498	AC107498 Rattus no
1000	20	47.6	253594	12	AC128893	AC128893 Rattus no

ALIGNMENTS

RESULT 1	BD233463	Human protein having hydrophobic domain and DNA encoding the same.	603 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD233463					
DEFINITION	BD233463					
ACCESSION	BD233463.1	GI:33043233				
VERSION	JP 2002519016-A/9.					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1	(bases 1 to 603)				
AUTHORS	Kato,S. and Kimura,T.					
TITLE	Human protein having hydrophobic domain and DNA encoding the same					
JOURNAL	Patent: JP 2002519016-A 9 02-JUL-2002;					
COMMENT	SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC					
OS	Homo sapiens (human)					
PN	JP 2002519016-A/9					
PD	02-JUL-2002					
PF	18-JUN-1999	JP 2000557267				
PI	SEISHI KATO,TOMOKO KIMURA					
PC	C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N5/00,C12N5/00					
CC	Human protein having hydrophobic domain and DNA encoding the					
CC	same					
FH	Key	Location/Qualifiers				
FT	source	1..603				
FT		/organism='Homo sapiens (human)'				
FEATURES	source	Location/Qualifiers				
		1..603				
		/organism='Homo sapiens'				
		/mol_type='genomic DNA'				
		/db_xref='taxon:9606'				

ORIGIN	Query Match	100.0%;	Score 42;	DB 2;	Length 603;	
	Best Local Similarity	100.0%;	Pred. No. 8.9e-08;			
	Matches	42;	Conservative	0;	Mismatches	0; Gaps 0;
QY	1	GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42				
DB	550	GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 591				
RESULT 2	BD209699	1212 bp	DNA	linear	PAT 17-JUL-2003	
LOCUS	BD209699					
DEFINITION	Compositions isolated from skin cells and methods for their use.					
ACCESSION	BD209699					
VERSION	BD209699.1	GI:33019469				
KEYWORDS	JP 2002512798-A/171.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1	(bases 1 to 1212)				
AUTHORS	Strachan,L., Sleeman,M., Watson,J.D., Onrust,R., Kumble,A. and Murison,J.G.					
TITLE	Compositions isolated from skin cells and methods for their use					
JOURNAL	Patent: JP 2002512798-A 171 08-MAY-2002;					
COMMENT	GENESIS RESEARCH AND DEVELOPMENT CORP LTD					
	OS Homo sapiens (human)					
	PN JP 2002512798-A/171					
	PD 08-MAY-2002					
	PF 29-APR-1999	JP 2000546009				
	PR 09/069726,09-NOV-1998	US 09/188930				
	LORNA STRACHAN,MATTHEW SLEEMAN,JAMES DOUGLAS WATSON,RENE PI ONRUST,					
PI	ANAND KUMBLE,JAMES GREG MURISON					
PC	C12N15/09,A61K38/00,A61P9/00,A61P17/00,A61P29/00,A61P31/18, PC A61P35/00,					
PC	C07K14/47,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12N15/00,					
CC	A61K37/02,C12N5/00					
CC	Compositions isolated from skin cells and methods for their					
CC	use.					
FH	Key	Location/Qualifiers				
FT	source	1..1212				
FT		/organism='Homo sapiens (human)'				
FEATURES	source	Location/Qualifiers				
		1..1212				
		/organism='Homo sapiens'				
		/mol_type='genomic DNA'				
		/db_xref='taxon:9606'				
ORIGIN	Query Match	100.0%;	Score 42;	DB 2;	Length 1212;	
	Best Local Similarity	100.0%;	Pred. No. 1.1e-07;			
	Matches	42;	Conservative	0;	Mismatches	0; Gaps 0;
QY	1	GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42				
DB	990	GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1031				
RESULT 3	AR341505	1212 bp	DNA	linear	PAT 17-AUG-2003	
LOCUS	AR341505					
DEFINITION	Sequence 249 from patent US 6573095.					
ACCESSION	AR341505					
VERSION	AR341505.1	GI:33733640				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					


```
Unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Strachan,L.
TITLE Polynucleotides isolated from skin cells
JOURNAL Patent: US 6573095-A 249 03-JUN-2003;
Genesis Research & Development Corporation Limited; Parnell;
NZX;
FEATURES
source Location/Qualifiers
1..1212
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 1212;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 990 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 1031
|||||

RESULT 4
AX078375 1619 bp DNA linear PAT 22-FEB-2001
LOCUS
DEFINITION Sequence 43 from Patent WO0107612.
ACCESSION AX078375
VERSION AX078375.1 GI:13158044
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
TITLE Receptors and associated proteins
JOURNAL Patent: WO 0107612-A 43 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..1619
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 2681738CB1"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 1137 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 1178
|||||

RESULT 5
BD233473 1718 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION BD233473
VERSION BD233473.1 GI:33043243
KEYWORDS
SOURCE JP 2002519016-A/19.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Kato,S. and Kimura,T.
TITLE Human protein having hydrophobic domain and DNA encoding the same
JOURNAL Patent: JP 2002519016-A 19 02-JUL-2002;

SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
OS Homo sapiens (human)
PN JP 2002519016-A/19
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO,TOMOKO KIMURA
PC
C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
FH Key same Location/Qualifiers
FT source 1..1718
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source Location/Qualifiers
1..1718
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 1718;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 561 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 602
|||||

RESULT 6
AK172760 1826 bp mRNA linear PRI 07-MAY-2004
LOCUS
DEFINITION Homo sapiens cDNA FLJ23921 fis, clone COL02043, highly similar to
Homo sapiens retinoic acid induced 3 (RAI3).
ACCESSION AK172760
VERSION AK172760.1 GI:47077732
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,K., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1826)
AUTHORS Sugano,S. and Suzuki,Y.
JOURNAL Direct Submission
Submitted (22-APR-2004) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:flcdna@mail.ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
source Location/Qualifiers
1..1826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL02043"
/tissue_type="colon"
/clone_lib="COL"
/note="Cloning vector pME18SFL3"
ORIGIN
```

```
Query Match      100.0%; Score 42; DB 5; Length 1826;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 638 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 679

RESULT 7
LOCUS CQ723177 2290 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 9111 from Patent WO02068579.
ACCESSION CQ723177
VERSION CQ723177.1 GI:42284034
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 9111 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..2290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 42; DB 2; Length 2290;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1117 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 8
LOCUS BC003665 2296 bp mRNA linear PRI 24-NOV-2004
DEFINITION Homo sapiens G protein-coupled receptor, family C, group 5, member
A, mRNA (CDNA clone MGC:923 IMAGE:2988011), complete cds.
ACCESSION BC003665
VERSION BC003665.2 GI:33872669
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Iqbal,M.N.A., Peters,G.J.,
Abramson,R.D., Mullen,L.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Vitaloni,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2296)
Director MGC Project.
Direct Submission
Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:13177795.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IMAGE Plate: 3 Row: 1 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 12056470.
FEATURES
Location/Qualifiers
1..2296
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:923 IMAGE:2988011"
/tissue type="Colon, adenocarcinoma"
/clone_lib="NIH MGC.15"
/lab_host="DH10B-R"
/note="vector: POTB7"
1..2296
/gene="GPCR5A"
/note="synonyms: RAIG1, GPCR5A, RA13"
/db_xref="GeneID:9052"
/db_xref="WIM:604138"
95..1168
/gene="GPCR5A"
/codon_start=1
/product="G protein-coupled receptor, family C, group 5,
member A"
/protein_id="AAH03665.1"
/db_xref="GI:13177796"
/db_xref="GeneID:9052"
/db_xref="WIM:604138"
/translation="MATTVPDGCNGLSKYKRLCDKAEAMGIVLETVATAGVVTSSVA
FMLTPTLVCKVQDSNNRRKMLPTQFLFLGVGLFTFAF1IGLDGSTGTPRFLFG
ILFSICSCLLAHAVSLTKVGRKPLSLVILGLAVGSFVADVIETVLTNRT
NNVVFSELSAPRNRNDFVLLITYVFLMALFFLMSSFTFCGFTGKRGHAIYLTML
LSIAIWWAITLMLPDFDRRDDTILLSALAANGWFLAYVSPFVLLTKORNPMD
YVPDFAFCFQLVKKSYGVENRAYSQBEITQGFETGDTLYAPYSTHFLQNLQNPQKE
FSIPRAHWAPSPYKDYEVKKEGS"
ORIGIN
Query Match      100.0%; Score 42; DB 5; Length 2296;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1112 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1153
```

RESULT 9
AF506289
LOCUS Homo sapiens orphan G protein-coupling receptor PEIG-1 mRNA, PRI 15-JUL-2002 complete cds.
ACCESSION AF506289
VERSION 1 GI:21779962
KEYWORDS Homo sapiens (human)
SOURCE ORGANISM
REFERENCE 1 (bases 1 to 2297)
AUTHORS Cafferata, E.G., Gonzalez-Guerrico, A.M., Costanzo, R., Pivetta, O.H. and Santa-Coloma, T.A.
TITLE Identification by differential display of a mRNA specifically induced by 12-O-tetradecanoylphorbol-13-acetate (TPA) in T84 human colon carcinoma cells
JOURNAL Cell. Mol. Biol. 42 (5), 797-804 (1996)
PUBMED 8832110
AUTHORS 2 (bases 1 to 2297)
REFERENCE Cafferata, E.G., Gonzalez-Guerrico, A.M., Costanzo, R., Pivetta, O.H. and Santa-Coloma, T.A.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2002) Laboratorio de Biologia Celular y Molecular, Instituto de Investigaciones Bioquimicas Fundacion Campomar, Patricias Argentinas 435, Buenos Aires 1405, Argentina
LOCATION/Qualifiers
FEATURES
1. .2297
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12.3"
/sex="male"
/cell_line="T84"
/tissue_type="colorectal carcinoma"
/note="derived from dbEST BE519991"
108. .1181
/note="belongs to the type 3 G protein-coupling receptor family; characterized by a 7-transmembrane domain motif; phorbol ester induced protein-1"
/codon_start=1
/product="orphan G protein-coupling receptor PEIG-1"
/protein_id="AAM77594.1"
/db_xref="GI:21779963"
/translation="MATTVPDGCNRGLKSKYYRLCDKAEAMGIVLETATAGVWTSVA FMLTPLIVCKVQDSNRKMLPTQFLFLGLVIGLTFGLTFAFIIGLDGSGTPRFLFG ILFSICFCLLAHAHVSILKLVGRKPLSLVILGLAVGFSLVQDVIAIEYIVLTWNRT NVNVSLSAPRNEDFVLLLTYYVLFMLATFLMSFTFCGSGFTGWRKHGAIHYLTML LSIATVAVITLMLPDRDWDITLSALAANGVFLIAYVSPFWLLTKQRNPM YPVEDAFCKPQLVKGYGVENRAYSQEETQGTFTGDTLVAPYSTHFLQVQPPQKE PSIPRAHAWSPYKDYEVKKEGS"
ORIGIN
Query Match 100.0%; Score 42; DB 5; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 1125 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1166
RESULT 10
AX930411
LOCUS Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA, PRI 29-DEC-1998 complete cds.
ACCESSION AF095448
VERSION 1 GI:4063889
KEYWORDS Homo sapiens (human)
SOURCE ORGANISM
REFERENCE 1 (bases 1 to 2302)
AUTHORS Cheng, Y. and Lotan, R.
TITLE Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G protein-coupled receptor
JOURNAL J. Biol. Chem. 273 (52), 35008-35015 (1998)
PUBMED 9857033
AUTHORS 2 (bases 1 to 2302)
REFERENCE Cheng, Y. and Lotan, R.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1998) Tumor Biology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA
LOCATION/Qualifiers
FEATURES
1. .2302
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p12.3-p13"
1. .2302
/gene="RAIG1"
/note="induced by retinoic acid"
100. .1173
/gene="RAIG1"
/codon_start=1
/product="putative G protein-coupled receptor"
/protein_id="AAC98506.1"
/db_xref="GI:4063889"
/translation="MATTVPDGCNRGLKSKYYRLCDKAEAMGIVLETATAGVWTSVA FMLTPLIVCKVQDSNRKMLPTQFLFLGLVIGLTFGLTFAFIIGLDGSGTPRFLFG ILFSICFCLLAHAHVSILKLVGRKPLSLVILGLAVGFSLVQDVIAIEYIVLTWNRT NVNVSLSAPRNEDFVLLLTYYVLFMLATFLMSFTFCGSGFTGWRKHGAIHYLTML

KEYWORDS Homo sapiens (human)
SOURCE ORGANISM
REFERENCE 1
AUTHORS Terrett, J.A.
TITLE Diagnosis of carcinoma using raigl polypeptides
JOURNAL Patent: WO 03087832-A 2 23-OCT-2003;
Oxford GlycoSciences (UK) Limited (GB)
LOCATION/Qualifiers
FEATURES
1. .2302
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 1117 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158
RESULT 11
AF095448
LOCUS Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA, PRI 29-DEC-1998 complete cds.
ACCESSION AF095448
VERSION 1 GI:4063889
KEYWORDS Homo sapiens (human)
SOURCE ORGANISM
REFERENCE 1 (bases 1 to 2302)
AUTHORS Cheng, Y. and Lotan, R.
TITLE Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G protein-coupled receptor
JOURNAL J. Biol. Chem. 273 (52), 35008-35015 (1998)
PUBMED 9857033
AUTHORS 2 (bases 1 to 2302)
REFERENCE Cheng, Y. and Lotan, R.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1998) Tumor Biology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA
LOCATION/Qualifiers
FEATURES
1. .2302
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p12.3-p13"
1. .2302
/gene="RAIG1"
/note="induced by retinoic acid"
100. .1173
/gene="RAIG1"
/codon_start=1
/product="putative G protein-coupled receptor"
/protein_id="AAC98506.1"
/db_xref="GI:4063889"
/translation="MATTVPDGCNRGLKSKYYRLCDKAEAMGIVLETATAGVWTSVA FMLTPLIVCKVQDSNRKMLPTQFLFLGLVIGLTFGLTFAFIIGLDGSGTPRFLFG ILFSICFCLLAHAHVSILKLVGRKPLSLVILGLAVGFSLVQDVIAIEYIVLTWNRT NVNVSLSAPRNEDFVLLLTYYVLFMLATFLMSFTFCGSGFTGWRKHGAIHYLTML

```

ORIGIN
LSIAIWAWITLLMLPDRDWDITLSSALAANGWFLLAYVSPFWLLTKQRPMD
IPVEDAFCKPOLVKSGYGVENRAYSQBEITQGFETGDTLYAPYSTHFQLQNPQPKQ
FSIPRAHAWPSPYKDYEVKKEGS"

Query Match      100.0%; Score 42; DB 5; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1117 GCCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 12
BD156680      2446 bp      DNA      linear      PAT 17-JAN-2003
LOCUS      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD156680
VERSION      BD156680.1 GI:27862438
KEYWORDS      JP 2002191363-A/11523.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1 (bases 1 to 2446)
AUTHORS      Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE      Primer for synthesizing full-length cDNA and use thereof
JOURNAL      Patent: JP 2002191363-A 11523 09-JUL-2002;
COMMENT      HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/11523
PD      09-JUL-2002
PE      28-JUL-2000 JP 2000280990
PF      TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI      SAITO,
PJ      JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PK      KEIICHI NAGAI, TETSUJI OTSUKI
PC      C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC      Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT      CDS (254)..(1324).
FEATURES      source
source      1..2446
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 42; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 13
AX877483      2446 bp      DNA      linear      PAT 17-DEC-2003
LOCUS      Sequence 12388 from Patent EP1074617.
ACCESSION      AX877483
VERSION      AX877483.1 GI:40032219
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1
AUTHORS      Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Saito, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Kikkawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, I., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T.O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsumura, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
```

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1
AUTHORS      Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE      Primers for synthesizing full-length cDNA and their use
JOURNAL      Patent: EP 1074617-A 12388 07-FEB-2001;
FEATURES      source
source      1..2446
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
254..1327
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89746.1"
/db_xref="GI:40032220"
/translation="MATTVPDCRCGLKSKYVRLCDKAEANGIVLETVATAGVWTSVA
FMLPLPLVCKVQDSNRKMLPTQFLFLLGLVGLFGLTLPAPLIGDGGTGPFRFLEFG
ILFICFCLLAHAVSLTKVGRKPLSLVLGLAVGSLVQDVIALEYIVLTWNRT
NNVFSLSAPRNEDEFVLLTYVLFMLTFLMSSFTFCGSPFKWKGHAHYITML
LSIAIWAWITLLMLPDRDWDITLSSALAANGWFLLAYVSPFWLLTKQRPMD
YPVEDAFCKPOLVKSGYGVENRAYSQBEITQGFETGDTLYAPYSTHFQLQNPQPKQ
FSIPRAHAWPSPYKDYEVKKEGS"

ORIGIN
Query Match      100.0%; Score 42; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 14
AK001761      2446 bp      mRNA      linear      PRI 20-JAN-2006
LOCUS      Homo sapiens CDNA FLJ10899 fis, clone NT2RP5003506.
ACCESSION      AK001761
VERSION      AK001761.1 GI:7023229
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1
AUTHORS      Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Saito, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Kikkawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, I., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T.O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsumura, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
```

Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Kanehori, K.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2446)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flij-cdna@nifty.com, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
Location/Qualifiers
1..2446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP5003506"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RPS"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
254..1327
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91890.1"
/db_xref="GI:7023230"
/translation="MATTVPDGRNGLSKYYRLCDKAEANGIVLETVATAGVVTSSA
FMLTLPILVCKVQDSNRKMLPTQFLFLGLVIGLTFATIGLDGSGTPTRPFLFG
ILFSICFSLLAHAYSLTKLVGRKPLSLVILGLVGFSLVQDVIALEYILVTNRT
NNVFSLSAPRNEFDVLLTYLVFLMALTFMSFTFCGFTGWRKHGHIYLTML
LSIAIWAIVITLLMLPDDRRWDDTILSSALAANGVFLLAYSPFVLLTKQRPMD
YVPEDAFKPQLVKSGYGVENRAYSQEITQGTFTGDTLYAPYSTHFLQLNQPPQKE
FSIPRAHWPSPKDYEVKKEGS"

Query Match 100.0%; Score 42; DB 5; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 15
LOCUS CQ981495 2456 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 350 from Patent EPI498424.
ACCESSION CQ981495
VERSION CQ981495.1 GI:58190785
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Bruemendorf, T., Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE Human nucleic acid sequences from lung tumours
JOURNAL Patent: EP 1498424-A 350 13-JAN-2005; Heiden, Esmeralda (DE); Hinemann, Bernd (DE); Hermann, Klaus (DE); Rosenthal, Andre (DE)

FEATURES
source
1..2456
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 16
LOCUS DD210040 2456 bp DNA linear PAT 19-JAN-2006
DEFINITION Screening for Modulators of Cancer.
ACCESSION DD210040
VERSION DD210040.1 GI:85654022
KEYWORDS JP 2005518782-A/34.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Zlotnik, A., Mack, D.H., Agiz, N., Gish, K.C., Hebeji, P.A., Wilson, K.E. and Afar, D.
TITLE Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer
JOURNAL Patent: JP 2005518782-A 34 30-JUN-2005; PROTEIN DESIGN LABS INC

COMMENT
OS Homo sapiens
PN JP 2005518782-A/34
PD 30-JUN-2005
PF 17-SEP-2002 JP 2003523912
PR 12-APR-2002 US 60/372246, 08-FEB-2002 US 60/355257, PR 08-FEB-2002 US 60/355145, 13-NOV-2001 US 60/350666, PR 20-SEP-2001 US 60/323887, 17-SEP-2001 US 60/323469 PI
Zlotnik, David H. Mack, Natasha Agiz, Kurt C Gish, Peter A Hebeji, PI Keith E Wilson, Daniel Afar

FEATURES
source
1..2456
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

```
RESULT 17
AX549168
LOCUS
DEFINITION Sequence 453 from Patent WO02061087.
ACCESSION AX549168
VERSION AX549168.1 GI:25813894
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS Burmer,G.C., Roush,C.L. and Brown,J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
JOURNAL Patent: WO 02061087-A 453 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
source
1..2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCACGGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
|||||

RESULT 18
AK122672
LOCUS
DEFINITION Homo sapiens CDNA FLJ16117 fis, clone ASTRO2003632, highly similar
to Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA.
ACCESSION AK122672
VERSION AK122672.1 GI:34527861
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Oabayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudoh,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Oho,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T.O., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kunagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,K.,
Tashiro,H., Tanigami,A., Fujiwara,T., Oho,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
```


JOURNAL Patent: WO 0142467-A 4043 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
source 1..3371
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 3371;
Best Local Similarity 100.0%; Pred. NO. 1.6e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1184 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1225
RESULT 20
CQ894732
LOCUS 6730 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 42 from Patent EP1471075.
ACCESSION CQ894732
VERSION CQ894732.1 GI:55467481
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Rosenthal,A.D., Pilarsky,C., Dahl,E., Specht,T., Bruemendorf,T.,
Lichtner,R., Staub,E., Roepcke,S. and Li,X.I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: EP 1471075-A 42 27-OCT-2004;
Hinzmann, Bernd (DE); Rosenthal, Andre (DE); Pilarsky, Christian
(DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie
(DE)
FEATURES Location/Qualifiers
source 1..6730
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 6730;
Best Local Similarity 100.0%; Pred. NO. 2.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1271 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
RESULT 21
AC007688/c
LOCUS 161577 bp DNA linear PRI 30-AUG-2002
DEFINITION Homo sapiens 12 BAC RPC111-392P7 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
ACCESSION AC007688
VERSION AC007688.15 GI:5815499
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 161577)
AUTHORS Muzny,D., Arenson,A.D., Bouck,J., Bunac,C., Chen,Z., Ding,Y.,
Dugan,S., Durbin,J., Forcum,J., Garcia,C., Gorrell,J.H.,
Gorrell,L.B., Hernandez,J., Issar,A., Jackson,L., Kneitz,S.,
Kondejewski,N., Lau,S., Leal,B., Lee,E., Lichtarge,O., Liu,W.,
Logan,O., Lu,J., Marondel,I., Martinez,C., Merscher,S., Miller,A.,
Montgomery,K., Oswal,G., Pampell,I.R., Parish,B.J., Perez,L.,
Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Shim,C., Simon,M.,
Vo,O., Williamson,A., Worley,K.C., Shang,A.M., Yang,R., Yu,W.,
Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 1, 1999 this sequence version replaced gi:5757565.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.
ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.
SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.
QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.
QUALSTAT-REPORT-----
----- Summary Statistics -----
Contig length: 161577
Phrap values in estimate: 160751

Average error rate (BCM-Phrap estimate): 0.000163691
Fraction of Phrap values less than 40 : 0.0376047
Number of consensus changing edits: 30
Number of N's in consensus : 0

Consensus changing edits	
Position	Original*Context Edited*Context
7033	actctgcgct (n) cgcgcccttt
47567	aaaaaataaa (n) ggaataataat
51135	aaagaataaa (n) aaagaataaa
75582	aaaaaataaa (n) aaagaataaa
75585	aaaaaataaa (n) ggaatgtcgc
75567	ctaaagcaga (n) taagatctta
75752	ttttaaatag (g) gctttgcttt
84017	ttttgttttt (n) tgtttttttt
85227	ttttgttttt (n) tgtttttttt
99681	ggagttcaag (n) atgcagtgag
111307	aattctttat (n) cccgaattca
111374	actagtatac (n) atcntttttt
111378	gtatacnatc (n) tttttttttt
113621	aaaaaataaa (n) ccatcctaga
135812	attcaccttc (c) tttttttttt
137207	tgcaggcac (n) cgcacacacg
137218	cgcacacacg (n) ctggttaagt
145113	gcaagtgaat (n) natgtagaat
145114	caagtgaan (n) atgtagaata
145232	ggacccggtg (n) ntacnccag
145233	ggacccggtg (n) tcacnccagt
145238	ggtgntcac (n) ccagtaatcc
145321	gcaacatggt (c) nmcccatcc
145322	caacatggtt (n) nmcccatct
145323	caacatggtt (n) nmcccatct
145324	acatggtttn (n) ccccatctct
145363	gttcaccagg (n) gtggtggcgt
145377	gtggtggcgt (n) gtggtggcct
145397	tgactacttg (n) gagggaggat
145505	aaaaaaaaaa (n) gaaaaaaaaa

----- Distribution of Quality < 40 Bases -----

# bases	Phrap Value Range															
	5	10	15	20	25	30	35	40								
1000	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
900	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
800	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
700	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
600	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
500	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
400	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
300	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
200	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
100	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
0	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Version: 1.01 qvfo.
Location/Qualifiers
1.161577
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-392P7"
complement(3..65)
/rpt_family="MER4B"
161..471
/rpt_family="AluX"
complement(607..702)
/rpt_family="MIR"

repeat_region
repeat_region
repeat_region

repeat_region 703..838
/rpt_family="FLAM C"
repeat_region complement(839..916)
/rpt_family="MIR"
repeat_region complement(1205..1377)
/rpt_family="MER104"
STS 1388..1525
/standard_name="WIAF-759-STS"
/db_xref="dbSTS:65358"
repeat_region 1683..1971
/rpt_family="AluJo"
repeat_region 1974..2261
/rpt_family="Aluub"
repeat_region 2486..2593
/rpt_family="MIR"
misc_feature 2610..3214
/note="Region: Unigene cluster similar to AA056332 and AI074576"
STS 2712..2823
/standard_name="SHGC-44583"
/db_xref="dbSTS:48451"
repeat_region 3185..3491
/rpt_family="AluSg"
repeat_region complement(6397..6537)
/rpt_family="L2"
repeat_region 6967..7158
/rpt_family="(CCCCG)n"
repeat_region 7094..7240
/rpt_family="(CCG)n"
repeat_region complement(7770..8250)
/rpt_family="MER44B"
repeat_region 8831..9073
/rpt_family="MER102"
repeat_region complement(9075..9303)
/rpt_family="L2"
repeat_region 9671..10135
/rpt_family="LTR33A"

Query Match 100.0%; Score 42; DB 5; Length 161577;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 94911 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 94870

RESULT 22
BD150346 680 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD150346
ACCESSION BD150346
VERSION BD150346.1 GI:27856104
KEYWORDS JP 2002191363-A/5189.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 5189 09-JUL-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/5189
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOgai,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI

```

PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1. 680
/organism='Homo sapiens (human)'
FEATURES
source
1. 680
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 80.0%; Score 33.6; DB 2; Length 680;
Best Local Similarity 87.8%; Pred. No. 0.0007;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAA 41
|||||
Db 633 GCCACGCTTGCCCAACCTTACAAAGACTATGAANGTAA 673

RESULT 23
AX870284
LOCUS AX870284 680 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 5189 from Patent EP1074617.
ACCESSION AX870284
VERSION AX870284.1 GI:40025147
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 5189 07-FEB-2001;
Research Association for Biotechnology (JP)
LOCATION/Qualifiers
FEATURES
source
1. 680
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 80.0%; Score 33.6; DB 2; Length 680;
Best Local Similarity 87.8%; Pred. No. 0.0007;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAA 41
|||||
Db 633 GCCACGCTTGCCCAACCTTACAAAGACTATGAANGTAA 673

RESULT 24
AC132640
LOCUS AC132640 194709 bp DNA linear HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-339B14, WORKING DRAFT SEQUENCE.
ACCESSION AC132640
VERSION AC132640.3 GI:25007393
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 194709)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

```

```

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Bunay, C., Burch, K., Burrell, K., Caldwell, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Dreher, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhewa, L., Louiseged, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montanayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 194709)
Rat Genome Sequencing Consortium.
Submitted (02-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 194709)
Rat Genome Sequencing Consortium.
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23908481.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature

```

Center: Baylor College of Medicine

REFERENCE

4 (bases 1 to 235468)
 Worley, K.C.
 Direct Submission
 Submitted (04-JUN-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 4, 2003 this sequence version replaced gi:24942020.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

AUTHORS
TITLE
JOURNAL

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
 standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

ml.

FEATURES

source Location/Qualifiers
 1. .235468
 /organism="Rattus norvegicus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10116"
 /chromosomes="3"
 /clone="CH230-239D4"
 complement(1..979)
 /rpt_family="RNHAL1"
 1043..1229
 /rpt_family="B2_Mm2"
 complement(1997..2128)
 /rpt_family="B1_Mm"
 2502..2584
 /rpt_family="B4A"
 2905..2997
 /rpt_family="MIR3"
 4150..4266
 /rpt_family="B1_Rn"
 4270..4290
 /rpt_family="AT_rich"
 4510..5152
 /rpt_family="RMER6A"
 5443..5578
 /rpt_family="ID_B1"
 complement(5470..5586)
 /rpt_family="YREP_Mm"
 5587..5683
 /rpt_family="ID_RN"

repeat_region complement(5684..5708)
 /rpt_family="YREP_Mm"
 5720..5914
 /rpt_family="B3"
 5924..6088
 /rpt_family="ID_B1"
 complement(7437..7533)
 /rpt_family="ID_Rn"
 7928..8091
 /rpt_family="MTD"
 8092..8392
 /rpt_family="MTD"
 8424..8571
 /rpt_family="MIRb"
 complement(8663..8889)
 /rpt_family="L2"
 complement(9086..9167)
 /rpt_family="L2"
 complement(9218..9313)
 /rpt_family="ID_Rn"
 complement(9380..9428)
 /rpt_family="ID_B1"
 complement(10423..10540)
 /rpt_family="PB1D9"
 complement(11286..11359)
 /rpt_family="ID_B1"
 complement(11360..11392)
 /rpt_family="B1F"
 complement(11497..11612)
 /rpt_family="RSINE1"
 11970..12086
 /rpt_family="PB1D10"
 12844..12985
 /rpt_family="B1_Mm"
 complement(13014..13122)
 /rpt_family="RSINE1"
 complement(13380..13470)
 /rpt_family="PB1D10"
 14058..14288
 /rpt_family="B4"
 14353..14544
 /rpt_family="B2_Mm2"
 15146..15323
 /rpt_family="B2_Rn"
 15324..15482
 /rpt_family="B2_Rn"
 15483..15516
 /rpt_family="TAAA)n"
 15517..15538
 /rpt_family="B2_Rn"
 complement(15532..15729)
 /rpt_family="B3"
 16345..16451
 /rpt_family="ID_Rn"
 16858..17020
 /rpt_family="B2_Mm2"
 complement(17339..17460)
 /rpt_family="RSINE1"
 complement(17467..17537)
 /rpt_family="BC1_Rn"

Query Match 58.6%; Score 24.6; DB 6; Length 235468;

Best Local Similarity 76.9%; Pred. No. 75;

Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 CCACGGTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41

Db 210151 CCATGGTGGCAAGCCCTGACAAAGACTAGGAATGAA 210189

RESULT 28

AC126155

LOCUS AC126155 253631 bp DNA linear HTG 13-MAY-2003

```

DEFINITION Rattus norvegicus clone CH230-7M5, WORKING DRAFT SEQUENCE.
ACCESSION AC126155
VERSION AC126155.4 GI:30579514
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 253631)
Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alebrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Royagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,B., Benahmed,F.,
Biswal,N., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,I., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunarathne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuwa,L., Louised,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Manghney,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Mitosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaekemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Fu,D.-L.,
Fuazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 253631)
Worley,K.C.
Direct Submission
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253631)
Rat Genome Sequencing Consortium.
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT
On May 13, 2003 this sequence version replaced gi:22856732.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKAR
Center clone name: CH230-7M5
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 245168 bases at least Q40
Consensus quality: 248824 bases at least Q30
Consensus quality: 247704 bases at least Q20
Estimated insert size: 254802; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 253631: contig of 253631 bp in length.
* Location/Qualifiers
* 1..253631
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-7M5"

FEATURES
source
Query Match 58.6%; Score 24.6; DB 12; Length 253631;
Best Local Similarity 76.9%; Pred. No. 77;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

ORIGIN
QY 3 CCACGCTTGGCGGCCCTTACAAAGACTATGAGCTAAA 41
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37759 CCATGGTGGCAAGCCCTGACAAAGACTAGGAATGAA 37797

RESULT 29
AC116169 189833 bp DNA linear HTG 27-MAR-2003
DEFINITION Homo sapiens clone RP11-55C13, 2 ordered pieces.
ACCESSION AC116169
VERSION AC116169.5 GI:23129007
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 189833)
Birren,B., Nusbaum,C. and Lander,E.
REFERENCE
AUTHORS

```

TITLE JOURNAL REFERENCE AUTHORS	Homo sapiens chromosome, clone RP11-55C13		* 113514 113613: gap of 100 bp	
	2 (bases 1 to 189833)		* 113614 189833: contig of 76220 bp in length.	
FEATURES source	Birren,B., Linton,J., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gordon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Rago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehocsky,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,K., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connell,T., O'Neill,D., O'Neill,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		Location/Qualifiers 1. .189833 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="RP11-55C13" /clone_lib="RP11-11 Human Male BAC" 113514..113613 /estimated_length=100	
gap	Query Match 58.1%; Score 24.4; DB 12; Length 189833; Best Local Similarity 73.8%; Pred No. 86; Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;		1. .189833 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="RP11-55C13" /clone_lib="RP11-11 Human Male BAC" 113514..113613 /estimated_length=100	
	ORIGIN		AC157012 229303 bp DNA linear HTG 01-JUL-2005 Bos taurus clone CH240-62M20, *** SEQUENCING IN PROGRESS ***, 27 unordered pieces.	
QY	1 GCCCAGCGTTGCCGCGCCCTTACAAAGACTATGAGTAAAG 42		AC157012.3 GI:68303464 HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.	
	Db 140251 GCCCAGCGCTGCTGAGCTCTCAGCAATTCATCGAGTAAAG 140292		Bos taurus (cattle)	
RESULT 30 AC157012 LOCUS	AC157012 229303 bp DNA linear HTG 01-JUL-2005 Bos taurus clone CH240-62M20, *** SEQUENCING IN PROGRESS ***, 27 unordered pieces.		Bos taurus	
	DEFINITION		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.	
ACCESSION VERSION KEYWORDS SOURCE ORGANISM	AC157012.3 GI:68303464 HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.		1 (bases 1 to 229303)	
	REFERENCE		Muzny,D., Maric,M., Metzker,M., Abruzzo,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Andra,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerri,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensheva,L., Louised,H., Lozada,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmood,M., Mallory,K., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelmech,O., Okwono,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poinexter,A., Popovic,D., Primus,E., Fu,L.-L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,	

Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 229303)

Worley, K.C.

Direct Submission

Submitted (07-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 229303)

Cow Genome Sequencing Consortium.

Direct Submission

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On June 29, 2005 this sequence version replaced gi:58743475.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FDOO

Center clone name: CH240-62M20

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 213349 bases at least Q40

Consensus quality: 215963 bases at least Q30

Consensus quality: 218476 bases at least Q20

Estimated insert size: 220573; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 9337: contig of 9337 bp in length

* 9338 10432: gap of 1095 bp

* 10433 13724: contig of 3292 bp in length

* 13725 17774: gap of 50 bp

* 13775 16987: contig of 3213 bp in length

* 16988 17087: gap of unknown length

* 17088 20724: contig of 3637 bp in length

* 20725 20774: gap of 50 bp

* 20775 26078: contig of 5304 bp in length

* 26079 26128: gap of 50 bp

* 26129 28649: contig of 2521 bp in length

* 28650 29630: gap of 981 bp

* 29631 34671: contig of 5040 bp in length

* 34671 34720: gap of 50 bp

* 34721 40011: contig of 5291 bp in length

* 40012 40111: gap of unknown length

* 40112 88972: contig of 48861 bp in length

* 88973 89022: gap of 50 bp

* 89023 111089: contig of 22067 bp in length

* 111090 111139: gap of 50 bp

* 11140 131191: contig of 20052 bp in length

* 131192 132275: gap of 1084 bp

* 132276 133460: contig of 1185 bp in length

* 133461 133560: gap of unknown length

* 133561 137385: contig of 3825 bp in length

* 137386 137435: gap of 50 bp

* 137436 144746: contig of 7311 bp in length

* 144747 144846: gap of unknown length

* 144847 145973: contig of 1127 bp in length

* 145974 146023: gap of 50 bp

* 146024 148479: contig of 2456 bp in length

* 148480 148579: gap of unknown length

* 148580 148669: contig of 1290 bp in length

* 149870 150090: gap of 221 bp

* 150091 160007: contig of 9917 bp in length

* 160008 160057: gap of 50 bp

* 160058 163938: contig of 3881 bp in length

* 163939 163988: gap of 50 bp

* 163989 177724: contig of 13736 bp in length

* 177725 177774: gap of 50 bp

* 177775 179811: contig of 2037 bp in length

* 179812 180606: gap of 795 bp

* 180607 194799: contig of 14193 bp in length

* 194800 194849: gap of 50 bp

* 194850 209924: contig of 15075 bp in length

* 209925 209974: gap of 50 bp

* 209975 215978: contig of 6004 bp in length

* 215979 216028: gap of 50 bp

* 216029 224039: contig of 8011 bp in length

* 224040 225433: gap of 1394 bp

* 225434 228175: contig of 2742 bp in length

* 228176 228275: gap of unknown length

* 228276 229303: contig of 1028 bp in length.

FEATURES

source

1. 229303
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-62M20"
9338. 10432
/estimated_length=1095
13725. 13774
/estimated_length=50
16988. 17087
/estimated_length=unknown
20725. 20774
/estimated_length=50
26079. 26128
/estimated_length=50
28650. 29630
/estimated_length=981
34671. 34720
/estimated_length=50
40012. 40111
/estimated_length=unknown
88973. 89022
/estimated_length=50
111090. 111139
/estimated_length=50
131192. 132275
/estimated_length=1084

Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, I., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (10-AUG-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 10, 2005 this sequence version replaced gi:62543423.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: J5628

Center clone name: 302_P_9

Some of the sequence contained within base pairs 1 - 61047 was stolen from accession AC102630
Some of the sequence contained within base pairs 75127 to the end of the clone was stolen from accession AC102564.

FEATURES

source

1..185146
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-302P9"
/clone_lib="RPCI-23 Female Mouse BAC"
1..72
/rpt_family="B1_Mur2"
1088..1438
/rpt_family="MTB"
1477..1583
/rpt_family="B1_Mus1"
1937..1984
/rpt_family="(TTTTTA)n"
complement(1986..2132)
/rpt_family="B1_Mus2"
complement(2592..2723)
/rpt_family="B1_Mur1"
2730..2753
/rpt_family="(T)n"
complement(2755..2895)
/rpt_family="B1_Mus2"
complement(3848..3974)
/rpt_family="B1_Mur3"
complement(4415..4490)
/rpt_family="LTR48"
complement(6248..6388)
/rpt_family="MER5B"
7656..7684
/rpt_family="(T)n"
7744..7799
/rpt_family="MIR"
8260..8288
/rpt_family="(CA)n"

repeat_region
complement(8323..9649)
/rpt_family="L1_Mus3"
repeat_region
10409..10519
/rpt_family="MTD"
repeat_region
10518..10617
/rpt_family="B1_Mur2"
11110..11504
/rpt_family="MTB_Mm"
repeat_region
11635..11789
/rpt_family="RSINE1"
repeat_region
11911..12075
/rpt_family="Lx3B"
complement(12727..12866)
/rpt_family="B1_Mur2"
complement(13372..13559)
/rpt_family="RMER15"
14231..14265
/rpt_family="(A)n"
complement(14284..14404)
/rpt_family="B1F2"
complement(14806..14900)
/rpt_family="PB1D7"
complement(15044..15250)
/rpt_family="B3"
15889..16181
/rpt_family="MT2_Mm"
16182..16890
/rpt_family="MERVL_Mm"
16886..21308
/rpt_family="MERVL_Mm"
21316..21402
/rpt_family="ORR1A3-int"
21403..21440
/rpt_family="MERVL_Mm"
21441..21933
/rpt_family="MT2_Mm"
23018..23181
/rpt_family="B4A"
23247..23413
/rpt_family="MTEa"
23712..23848
/rpt_family="B1_Mur3"
23916..24524
/rpt_family="RMERIC"
complement(24525..24836)
/rpt_family="ORR1A4"
24837..24886
/rpt_family="RMERIC"
complement(24899..25068)
/rpt_family="Lx7"
25075..25108
/rpt_family="(CA)n"
complement(25115..25612)
/rpt_family="Lx7"
26642..26722
/rpt_family="RSINE1"
26808..26828
/rpt_family="GC_rich"
26812..26832
/rpt_family="GC_rich"
complement(27289..27365)
/rpt_family="MLTL"
27479..27563
/rpt_family="(GA)n"
27618..27771
/rpt_family="RSINE1"
27842..28036
/rpt_family="ID_B1"
28054..28108
/rpt_family="ID4"
complement(28312..28635)
/rpt_family="MTE2b"
complement(28626..28706)
repeat_region

[illegible]

```
repeat_region 1468..1491
/rpt_family="AT_rich"
repeat_region complement(2746..2827)
/rpt_family="MIR"
repeat_region 3070..3313
/rpt_family="L2"
repeat_region complement(3536..3591)
/rpt_family="MIR"
repeat_region 3985..4358
/rpt_family="ORR1D"
repeat_region 5146..5265
/rpt_family="RSINE1"
repeat_region 5266..5297
/rpt_family=" (CTCA)n"
repeat_region 5308..5353
/rpt_family=" (CAGA)n"
repeat_region 5537..5727
/rpt_family="B2_Mm1"
repeat_region 5731..5877
/rpt_family="B1_MM"
repeat_region 5880..5949
/rpt_family="GA-rich"
repeat_region 6044..6181
/rpt_family="B1_MM"
repeat_region 6182..6218
/rpt_family=" (CAAAA)n"
repeat_region 7488..7586
/rpt_family="L2"
repeat_region complement(7628..7722)
/rpt_family="MIR"
repeat_region 8385..8541
/rpt_family="B3A"
repeat_region 8548..8590
/rpt_family=" (TCTG)n"
repeat_region 8604..8673
/rpt_family="B1F"
repeat_region 8674..8731
/rpt_family=" (GA)n"
repeat_region 8732..8762
/rpt_family="ID_B1"
repeat_region 9094..9219
/rpt_family="MER34C"
repeat_region complement(9374..9492)
/rpt_family="PB1"
repeat_region complement(10096..10217)
/rpt_family="Lx6"
repeat_region 10498..10747
/rpt_family="B4A"
repeat_region 10748..10870
/rpt_family="RSINE1"
repeat_region 10897..110923
/rpt_family=" (CAAA)n"
repeat_region 10979..11028
/rpt_family=" (CATA)n"
repeat_region 11079..11154
/rpt_family="B1F"
repeat_region 11155..11174
/rpt_family="ID_B1"
repeat_region 11241..11311
/rpt_family=" (TTTC)n"
repeat_region complement(11313..11471)
/rpt_family="B1_MM"
repeat_region 12420..12442
/rpt_family=" (TTTA)n"
repeat_region 12442..12509
/rpt_family=" (TG)n"
repeat_region complement(12519..12803)
/rpt_family="B4A"
repeat_region 12863..12911
/rpt_family=" (TA)n"
repeat_region 13046..13099
/rpt_family="B1_MM"
repeat_region 13106..13263
```

```
/rpt_family="GA-rich"
Query Match 54.8%; Score 23; DB 6; Length 230474;
Best Local Similarity 83.9%; Pred. No. 4.1e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCGAGCCCTTACAAAGACTAT 33
DB 94133 CCACCTTGGCGAGCCCTTGCCAAAGTTAT 94163

RESULT 36
AR630602 LOCUS 1500 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 3 from patent US 6841374.
ACCESSION AR630602
VERSION AR630602.1 GI:59765853
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Ishikawa,T., Yoshi,H., Onishi,T., Imagawa,T. and Ishibaishi,M.
TITLE Enhanced immunogen for inactivated vaccine for infection with
Japanese encephalitis viruses and process for producing the same
JOURNAL Patent: US 6841374-A 3 11-JAN-2005;
Research Foundation for Microbial Diseases of Osaka University;
Osaka;
JPX;

FEATURES             Location/Qualifiers
     source           1..1500
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
Query Match 54.3%; Score 22.8; DB 2; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
DB 1203 CACGCTGGCAAGCCCTTTCAACGACTTGAAG 1236

RESULT 37
AF112297 LOCUS 1500 bp RNA linear VRL 12-APR-1999
DEFINITION Japanese encephalitis virus (STRAIN NAKAYAMA) envelope protein
ACCESSION AF112297
VERSION AF112297.1 GI:4580915
KEYWORDS
SOURCE
ORGANISM Japanese encephalitis virus strain Nakayama
Japanese encephalitis virus strain Nakayama
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Chin,C., Sung,S.-C. and Chang,G.-J.
TITLE Genetic variation of envelope protein gene of Japanese encephalitis
virus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1500)
AUTHORS Chin,C., Sung,S.-C. and Chang,G.-J.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1998) Division of Vector-Borne Infectious
Diseases, Centers for Disease Control and Prevention, P.O. Box
2087, Fort Collins, CO 80522, USA

FEATURES             Location/Qualifiers
     source           1..1500
                     /organism="Japanese encephalitis virus strain Nakayama"
                     /mol_type="genomic RNA"
                     /strain="Nakayama"
                     /db_xref="taxon:11076"
```

```

/notes="from Walter Reed Army Institute of Research"
<1..>1500
/codon_start=1
/product="envelope protein"
/protein_id="AAD24538.1"
/db_xref="GI:4580916"
/translation="FNCILGMGNRDFIEGASGATWVDLVLEGDSCLTLMANDKPTLDVR
MINIEAVQLAEVRSYCYHASVTDISTVARCPPTGEAHNEKRADSSVYVCKQGFTRDGMG
NGCGLFGKSGIDTCAKFSCSTKAIGRTIQPENIKYEVGIFVHGTTTSENHNGYSAQVG
ASQAAKFTVTNAPSITLKLGDYGEVTLDCPRSLNTEAFYVMTVSGSKFLVHREWF
HDALPWTSPSTAWNRRELLMESEAHATKQSVVALGSOEGGLHQALAGAIIVVEYSN
SVKLTSGLHKLKMDLKAIGTTCYGMCTEKFSAKNPADTGHGTWVIELSYGSDGP
CKIPIVSVASLNDMPVGLVTVNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWKAGSTLTKAFSTTLKGAORLAALGDTAWDFGSGGVFNSIGKAVHQVFGGAFR
TLFGGMSWITQGLMGALLLWMGVNARDRSIALAFLATGGLVFLATNVHA"

ORIGIN
Query Match      54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
    ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGGCAAGCCTTTTCAACGACTTTGAAG 1236

RESULT 39
AY243826 Japanese encephalitis virus isolate Tla polyprotein gene, partial
LOCUS AY243826 1500 bp RNA linear VRL 27-APR-2003
DEFINITION cds.
ACCESSION AY243826
VERSION AY243826.1 GI:30143712
KEYWORDS
SOURCE
ORGANISM Japanese encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Liang, G., Li, X. and Fu, S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2003) State Key Lab. of Molecular Virology &
Genetic Engineering, Inst. of Virology, 100 Yingxin St., Xuan Wu
Qu, Beijing 100052, China
FEATURES
source
1..1500
/organism="Japanese encephalitis virus"
/mol_type="genomic RNA"
/isolate="Tla"
/isolation_source="patient brain sample"
/db_xref="taxon:11072"
/country="China"
<1..>1500
/codon_start=1
/product="polyprotein"
/protein_id="AAP14877.1"
/db_xref="GI:30143713"
CDS
translation="FNCILGMGNRDFIEGASGATWVDLVLEGDSCLTLMANDKPTLDVR
MINIEAVQLAEVRSYCYHASVTDISTVARCPPTGEAHNEKRADSSVYVCKQGFTRDGMG
NGCGLFGKSGIDTCAKFSCSTKAIGRTIQPENIKYEVGIFVHGTTTSENHNGYSAQVG
ASQAAKFTVTNAPSITLKLGDYGEVTLDCPRSLNTEAFYVMTVSGSKFLVHREWF
HDALPWTSPSTAWNRRELLMESEAHATKQSVVALGSOEGGLHQALAGAIIVVEYSN
SVKLTSGLHKLKMDLKAIGTTCYGMCTEKFSAKNPADTGHGTWVIELSYGSDGP
CKIPIVSVASLNDMPVGLVTVNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWKAGSTLTKAFSTTLKGAORLAALGDTAWDFGSGGVFNSIGKAVHQVFGGAFR
TLFGGMSWITQGLMGALLLWMGVNARDRSIALAFLATGGLVFLATNVHA"

ORIGIN
Query Match      54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
    ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGGCAAGCCTTTTCAACGACTTTGAAG 1236

RESULT 40
AY243843 Japanese encephalitis virus isolate A2 polyprotein gene, partial
LOCUS AY243843 1500 bp RNA linear VRL 27-APR-2003
DEFINITION cds.
ACCESSION AY243843
VERSION AY243843.1 GI:30143745
KEYWORDS
SOURCE
ORGANISM Japanese encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)

```

```

/notes="from Walter Reed Army Institute of Research"
<1..>1500
/codon_start=1
/product="envelope protein"
/protein_id="AAD24538.1"
/db_xref="GI:4580916"
/translation="FNCILGMGNRDFIEGASGATWVDLVLEGDSCLTLMANDKPTLDVR
MINIEAVQLAEVRSYCYHASVTDISTVARCPPTGEAHNEKRADSSVYVCKQGFTRDGMG
NGCGLFGKSGIDTCAKFSCSTKAIGRTIQPENIKYEVGIFVHGTTTSENHNGYSAQVG
ASQAAKFTVTNAPSITLKLGDYGEVTLDCPRSLNTEAFYVMTVSGSKFLVHREWF
HDALPWTSPSTAWNRRELLMESEAHATKQSVVALGSOEGGLHQALAGAIIVVEYSN
SVKLTSGLHKLKMDLKAIGTTCYGMCTEKFSAKNPADTGHGTWVIELSYGSDGP
CKIPIVSVASLNDMPVGLVTVNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWKAGSTLTKAFSTTLKGAORLAALGDTAWDFGSGGVFNSIGKAVHQVFGGAFR
TLFGGMSWITQGLMGALLLWMGVNARDRSIALAFLATGGLVFLATNVHA"

ORIGIN
Query Match      54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
    ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGGCAAGCCTTTTCAACGACTTTGAAG 1236

RESULT 40
AY243843 Japanese encephalitis virus isolate A2 polyprotein gene, partial
LOCUS AY243843 1500 bp RNA linear VRL 27-APR-2003
DEFINITION cds.
ACCESSION AY243843
VERSION AY243843.1 GI:30143745
KEYWORDS
SOURCE
ORGANISM Japanese encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)

```

```

AUTHORS Liang,G., Li,X. and Fu,S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2003) State Key Lab. of Molecular Virology & Genetic Engineering, Inst. of Virology, 100 Yingxin St., Xuan Wu Qu, Beijing 100052, China
FEATURES
source
1. .1500
/organism="Japanese encephalitis virus"
/mol_type="genomic RNA"
/isolate="A2"
/isolation_source="patient brain sample"
/db_xref="taxon:11072"
/country="China"
<1..>1500
/notes="envelope protein region"
/codon_start=1
/product="polyprotein"
/protein_id="AAPI4893.1"
/db_xref="GI:30143746"
/translation="FNCIMGNRDFTIEGASGATWVDLVLEGSCLTIMANDKPTLDVR
MINIEAQLAEVRSYCYHASVTDISTVARCPTTGEAHNEKRDSSVVCQGTDRMG
NGCGLPGKGSIDTCAKFSCTKAIGRTIOENIKYEVGI FVHGTTTSENHGNYSAOVG
ASQAAKFTVTPNAPSITLKLGDYGEVTLDCPEPSGLNTEAFYVMTVGSKSLVHREWF
HDALPWTTPSSTAWNRRELLMFEFEAHATKQSVVALGSGEGGLHQAALAGAI VVEYSN
SVKLTSGLKRLKMDKLKGTTYMCTEKFSAKNPADTGHGTVVI ELSYSGSDGP
CKTPIVSVASLNDMTPVGRVLTVPNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWYKAGSTLKGAFSTTLKGAORLAALGDTAWDFGSI GGVFNSIGKAVHVFEGGAFR
TLFGGMSWITQGLMGALLLWGMVGNARDRSIALAFLATGGVLVFLATNVHA"

Query Match 54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

CDS
<1..>1500
/codon_start=1
/product="polyprotein"
/protein_id="AAPI4893.1"
/db_xref="GI:30143746"
/translation="FNCIMGNRDFTIEGASGATWVDLVLEGSCLTIMANDKPTLDVR
MINIEAQLAEVRSYCYHASVTDISTVARCPTTGEAHNEKRDSSVVCQGTDRMG
NGCGLPGKGSIDTCAKFSCTKAIGRTIOENIKYEVGI FVHGTTTSENHGNYSAOVG
ASQAAKFTVTPNAPSITLKLGDYGEVTLDCPEPSGLNTEAFYVMTVGSKSLVHREWF
HDALPWTTPSSTAWNRRELLMFEFEAHATKQSVVALGSGEGGLHQAALAGAI VVEYSN
SVKLTSGLKRLKMDKLKGTTYMCTEKFSAKNPADTGHGTVVI ELSYSGSDGP
CKTPIVSVASLNDMTPVGRVLTVPNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWYKAGSTLKGAFSTTLKGAORLAALGDTAWDFGSI GGVFNSIGKAVHVFEGGAFR
TLFGGMSWITQGLMGALLLWGMVGNARDRSIALAFLATGGVLVFLATNVHA"

ORIGIN
Query Match 54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CACGCTGGCGAGCCCTTACAAAGACTATGAAG 37
||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGCGAAGCCCTTTTCAACGACTTTGAAG 1236

RESULT 41
JEU44974 LOCUS 1500 bp RNA linear VRL 01-AUG-1996
DEFINITION Japanese encephalitis virus TL envelope protein gene, partial cds.
ACCESSION U44974
VERSION U44974.1 GI:1477521
KEYWORDS Japanese encephalitis virus
ORGANISM Japanese encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Sung,S.
TITLE Genetic variations of envelope protein gene of Japanese encephalitis viruses from Taiwan
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1500)
AUTHORS Sung,S.
TITLE Direct Submission
JOURNAL Submitted (02-JAN-1996) Shu-Chiun Sung, Division of Molecular Virology, National Defense Medical Center, Taipei, Taiwan, Republic of China
FEATURES
source
1. .1500
/organism="Japanese encephalitis virus"
/mol_type="genomic RNA"
/strain="TL"
/db_xref="taxon:11072"
/lab_host="Aedes albopictus"
<1..>1500
/function="structural protein"
/codon_start=1
/product="envelope protein"
/protein_id="AAB05394.1"

CDS
<1..>1500
/codon_start=1
/product="envelope protein"
/protein_id="AAB05394.1"

ORIGIN
Query Match 54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CACGCTGGCGAGCCCTTACAAAGACTATGAAG 37
||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGCGAAGCCCTTTTCAACGACTTTGAAG 1236

RESULT 42
JEU70389 LOCUS 1500 bp mRNA linear VRL 14-JAN-1997
DEFINITION Japanese encephalitis virus envelope protein mRNA, partial cds.
ACCESSION U70389
VERSION U70389.1 GI:1778225
KEYWORDS Japanese encephalitis virus
SOURCE Japanese encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Chin,C.C., Sung,S.C., Shih,C.M. and Chang,G.J.
TITLE Correlation of evolution and phenotype of Japanese encephalitis viruses
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1500)
AUTHORS Chang,G.J.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-1996) DVBID/NCID/CDC/US Department of Health and Human Services, Arbovirus Disease Branch, P.O. Box 2087, Fort Collins, CO 80522, USA
FEATURES
source
1. .1500
/organism="Japanese encephalitis virus"
/mol_type="mRNA"
/strain="Beijing 1/China/1949/Human"
/isolates="substrain from DVBID/NCID/CDC"
/db_xref="taxon:11072"
<1..>1500
/function="structural protein"
/codon_start=1
/product="envelope protein"
/protein_id="AAB40664.1"
/db_xref="GI:1778226"
/translation="FNCIMGNRDFTIEGASGATWVDLVLEGSCLTIMANDKPTLDVR
MINIEAQLAEVRSYCYHASVTDISTVARCPTTGEAHNEKRDSSVVCQGTDRMG
NGCGLPGKGSIDTCAKFSCTKAIGRTIOENIKYEVGI FVHGTTTSENHGNYSAOVG
ASQAAKFTVTPNAPSITLKLGDYGEVTLDCPEPSGLNTEAFYVMTVGSKSLVHREWF
HDALPWTTPSSTAWNRRELLMFEFEAHATKQSVVALGSGEGGLHQAALAGAI VVEYSN
SVKLTSGLKRLKMDKLKGTTYMCTEKFSAKNPADTGHGTVVI ELSYSGSDGP
CKTPIVSVASLNDMTPVGRVLTVPNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWYKAGSTLKGAFSTTLKGAORLAALGDTAWDFGSI GGVFNSIGKAVHVFEGGAFR
TLFGGMSWITQGLMGALLLWGMVGNARDRSIALAFLATGGVLVFLATNVHA"

Query Match 54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CACGCTGGCGAGCCCTTACAAAGACTATGAAG 37
||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGCGAAGCCCTTTTCAACGACTTTGAAG 1236

RESULT 43
JEU70389 LOCUS 1500 bp mRNA linear VRL 14-JAN-1997
DEFINITION Japanese encephalitis virus envelope protein mRNA, partial cds.
ACCESSION U70389
VERSION U70389.1 GI:1778225
KEYWORDS Japanese encephalitis virus
SOURCE Japanese encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Chin,C.C., Sung,S.C., Shih,C.M. and Chang,G.J.
TITLE Correlation of evolution and phenotype of Japanese encephalitis viruses
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1500)
AUTHORS Chang,G.J.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-1996) DVBID/NCID/CDC/US Department of Health and Human Services, Arbovirus Disease Branch, P.O. Box 2087, Fort Collins, CO 80522, USA
FEATURES
source
1. .1500
/organism="Japanese encephalitis virus"
/mol_type="mRNA"
/strain="Beijing 1/China/1949/Human"
/isolates="substrain from DVBID/NCID/CDC"
/db_xref="taxon:11072"
<1..>1500
/function="structural protein"
/codon_start=1
/product="envelope protein"
/protein_id="AAB40664.1"
/db_xref="GI:1778226"
/translation="FNCIMGNRDFTIEGASGATWVDLVLEGSCLTIMANDKPTLDVR
MINIEAQLAEVRSYCYHASVTDISTVARCPTTGEAHNEKRDSSVVCQGTDRMG
NGCGLPGKGSIDTCAKFSCTKAIGRTIOENIKYEVGI FVHGTTTSENHGNYSAOVG
ASQAAKFTVTPNAPSITLKLGDYGEVTLDCPEPSGLNTEAFYVMTVGSKSLVHREWF
HDALPWTTPSSTAWNRRELLMFEFEAHATKQSVVALGSGEGGLHQAALAGAI VVEYSN
SVKLTSGLKRLKMDKLKGTTYMCTEKFSAKNPADTGHGTVVI ELSYSGSDGP
CKTPIVSVASLNDMTPVGRVLTVPNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWYKAGSTLKGAFSTTLKGAORLAALGDTAWDFGSI GGVFNSIGKAVHVFEGGAFR
TLFGGMSWITQGLMGALLLWGMVGNARDRSIALAFLATGGVLVFLATNVHA"

Query Match 54.3%; Score 22.8; DB 10; Length 1500;
Best Local Similarity 79.4%; Pred. No. 88;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CACGCTGGCGAGCCCTTACAAAGACTATGAAG 37
||||| ||| ||||| ||||| ||||| ||||| |||||
Db 1203 CACGCTGGCGAAGCCCTTTTCAACGACTTTGAAG 1236

RESULT 44
JEU70389 LOCUS 1500 bp mRNA linear VRL 14-JAN-1997
DEFINITION Japanese encephalitis virus envelope protein mRNA, partial cds.
ACCESSION U70389
VERSION U70389.1 GI:1778225
KEYWORDS Japanese encephalitis virus
SOURCE Japanese encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Chin,C.C., Sung,S.C., Shih,C.M. and Chang,G.J.
TITLE Correlation of evolution and phenotype of Japanese encephalitis viruses
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1500)
AUTHORS Chang,G.J.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-1996) DVBID/NCID/CDC/US Department of Health and Human Services, Arbovirus Disease Branch, P.O. Box 2087, Fort Collins, CO 80522, USA
FEATURES
source
1. .1500
/organism="Japanese encephalitis virus"
/mol_type="mRNA"
/strain="Beijing 1/China/1949/Human"
/isolates="substrain from DVBID/NCID/CDC"
/db_xref="taxon:11072"
<1..>1500
/function="structural protein"
/codon_start=1
/product="envelope protein"
/protein_id="AAB40664.1"
/db_xref="GI:1778226"
/translation="FNCIMGNRDFTIEGASGATWVDLVLEGSCLTIMANDKPTLDVR
MINIEAQLAEVRSYCYHASVTDISTVARCPTTGEAHNEKRDSSVVCQGTDRMG
NGCGLPGKGSIDTCAKFSCTKAIGRTIOENIKYEVGI FVHGTTTSENHGNYSAOVG
ASQAAKFTVTPNAPSITLKLGDYGEVTLDCPEPSGLNTEAFYVMTVGSKSLVHREWF
HDALPWTTPSSTAWNRRELLMFEFEAHATKQSVVALGSGEGGLHQAALAGAI VVEYSN
SVKLTSGLKRLKMDKLKGTTYMCTEKFSAKNPADTGHGTVVI ELSYSGSDGP
CKTPIVSVASLNDMTPVGRVLTVPNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI
NHHWYKAGSTLKGAFSTTLKGAORLAALGDTAWDFGSI GGVFNSIGKAVHVFEGGAFR
TLFGGMSWITQGLMGALLLWGMVGNARDRSIALAFLATGGVLVFLATNVHA"

Query Match 54.3%; Score 22.8; DB 1
```


Db	1203	CACGCTGGCGAAAGCCTTTTCAACGACTTTGAAG	1236
RESULT 43			
LOCUS	JEU70413	1500 bp	linear mRNA
DEFINITION	Japanese encephalitis virus envelope protein mRNA, partial cds.		VRL 14-JAN-1997
ACCESSION	U70413		
VERSION	U70413.1	GI:1778273	
KEYWORDS	Japanese encephalitis virus		
SOURCE	Japanese encephalitis virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Flavivirus; Japanese encephalitis virus group.		
REFERENCE	1 (bases 1 to 1500)		
AUTHORS	Chin, C.C., Sung, S.C., Shih, C.M. and Chang, G.J.		
TITLE	Correlation of evolution and phenotype of Japanese encephalitis viruses		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1500)		
AUTHORS	Chang, G.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-SEP-1996) DVBID/NCID/CDC/US Department of Health and Human Services, Arbovirus Disease Branch, P.O. Box 2087, Fort Collins, CO 80522, USA		
FEATURES	Location/Qualifiers		
source	1..1500		
	/organism="Japanese encephalitis virus"		
	/mol_type="mRNA"		
	/strain="Nakayama-NIH/1935/Japan/Human/brain"		
	/isolate="substrain from DVBID/NCID/CDC"		
	/db_xref="taxon:11072"		
	<1..>1500		
	/function="structural protein"		
	/codon_start=1		
	/product="envelope protein"		
	/protein_id="AAB40688.1"		
	/db_xref="GI:1778274"		
	/translation="FNLGMRDFTIEGASGATWVDLVLEGSCLTLMANDKPTLDVR MINIEAQLAEVRSYCHASVTDIETVARCPTTGEAHNEKRDSSVCKOGFTDRGNG NCGLFGKGSIDTCAKFSCTKAIGRTIQPENIKYEVGFVHGTTTSENHGNYSQVQ ASQAQFTVPNAPSITLKLGYGEVTLDCPRSGLNTFAFYMTVGSKFLVHREWF HDALPWTSPSTANRNRELLMESEAHATKOSAVALGSOEGGLHQAALAGAIIVVEYS SVKLTSGLHKLRLMDKLAKGTTCMCCTEKSFAKNPADTGHGTVTIVELSGSDGP CKIPIVSVASLNDMTVPGLVTVPNPFVATSSANSKVLVEMEPFGDSYIVVGRGDKQI NHHWKAGSTLGAFTTLKGAORLALGDTAWDFSGIGGVFNISGKAVHQVFGGAFR TLFGGMSWITQGLMGALLMGMVGNARDRSIALAFLATGGVLVFLATNVHA"		
CDS			
Query Match	54.3%	Score 22.8;	DB 10; Length 1500;
Best Local Similarity	79.4%	Pred. No. 88;	
Matches	27;	Conservative	0; Mismatches 7; Indels 0; Gaps 0;
QY	4	CACGCTTGGCGAGCCCTTACAAGACTATGAAG	37
Db	1203	CACGCTGGCGAAAGCCTTTTCAACGACTTTGAAG	1236
RESULT 45			
LOCUS	JEU70420	1500 bp	mRNA
DEFINITION	Japanese encephalitis virus envelope protein mRNA, partial cds.		VRL 14-JAN-1997
ACCESSION	U70420		
VERSION	U70420.1	GI:1778287	
KEYWORDS	Japanese encephalitis virus		
SOURCE	Japanese encephalitis virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Flavivirus; Japanese encephalitis virus group.		
REFERENCE	1 (bases 1 to 1500)		
AUTHORS	Chin, C.C., Sung, S.C., Shih, C.M. and Chang, G.J.		
TITLE	Correlation of evolution and phenotype of Japanese encephalitis viruses		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1500)		
AUTHORS	Chang, G.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-SEP-1996) DVBID/NCID/CDC/US Department of Health and Human Services, Arbovirus Disease Branch, P.O. Box 2087, Fort Collins, CO 80522, USA		
FEATURES	Location/Qualifiers		
source	1..1500		
	/organism="Japanese encephalitis virus"		
	/mol_type="mRNA"		
	/strain="VN118/1979/Vietnam/Mosquito"		
	/db_xref="taxon:11072"		
	<1..>1500		
	/function="structural protein"		
	/codon_start=1		
	/product="envelope protein"		
	/protein_id="AAB40695.1"		
	/db_xref="GI:1778288"		
CDS			
Query Match	54.3%	Score 22.8;	DB 10; Length 1500;
Best Local Similarity	79.4%	Pred. No. 88;	
Matches	27;	Conservative	0; Mismatches 7; Indels 0; Gaps 0;
QY	4	CACGCTTGGCGAGCCCTTACAAGACTATGAAG	37
Db	1203	CACGCTGGCGAAAGCCTTTTCAACGACTTTGAAG	1236
RESULT 44			
LOCUS	JEU70419	1500 bp	linear mRNA
DEFINITION	Japanese encephalitis virus envelope protein mRNA, partial cds.		VRL 14-JAN-1997
ACCESSION	U70419		
VERSION	U70419.1	GI:1778285	
KEYWORDS	Japanese encephalitis virus		
SOURCE	Japanese encephalitis virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Flavivirus; Japanese encephalitis virus group.		
REFERENCE	1 (bases 1 to 1500)		
AUTHORS	Chin, C.C., Sung, S.C., Shih, C.M. and Chang, G.J.		
TITLE	Correlation of evolution and phenotype of Japanese encephalitis viruses		

Beland, J., Chun, E., Del Rio, L., Gibson, S., Hanson, R., Kirkpatrick, R., Liu, J., Matsuo, C., Mayo, M., Santos, R. R., Stott, J., Tsai, M., Wong, D., Siddiqui, A., Holt, R., Jones, S. J. and Marra, M. A.
Direct Submission
Submitted (15-SEP-2005) BC Cancer Agency, Canada's Michael Smith Genome Sciences Centre, Suite 100, 570 West 7th Avenue, Vancouver, British Columbia V5Z 4S6, Canada
NTH-MGC Project
Contact: Robert Kirkpatrick
Canada's Michael Smith Genome Sciences Centre
BC Cancer Agency
Suite 100, 570 West 7th Avenue, Vancouver, British Columbia, Canada, V5Z 4S6
Tel: 1-604-707-5900 x5406
Fax: 1-604-876-3561
Email: robertk@bcsc.ca
Tissue Procurement: M. Taniguchi, Y. Meng, L. Guan and S. Moore, University of Alberta
cDNA Library Preparation: M. Masaaki, Y. Meng, S. Lobo, L. Guan and Dr. S. Moore
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Bovine Genome Sequencing Program, Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Moore S, Alexander L, Brownstein M, Guan L, Lobo S, Meng Y, Taniguchi M, Wang Z, Prange C, Schreiber K, Shennen C, Wagner L, Ali J, Bala M, Barber S, Babakoff R, Beland J, Chand S, Chun E, Del Rio L, Gibson S, Kirkpatrick R, Liu J, Matsuo C, Mayo M, Palmquist D, Santos RR, Stott J, Tsai M, Wong D, Wynhoven B, Hanson R, Siddiqui A, Holt R, Jones SJ, Marra MA

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 240 Row: n Column: 22.

FEATURES

Location/Qualifiers
1..2069
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Crossbred x Angus"
/db_xref="taxon:9913"
/clone="MGC:127231 IMAGE:7945740"
/tissue_type="ileum, crossbred x Angus female calf, 6 months old."
/clone_lib="CC_BGC-02"
/lab_host="XL10-Gold"
/notes="Vector: pCMV-SPORT6.0"
1..2069
/genes="MGC127231"
/db_xref="GeneID:516026"
136..1209
/genes="MGC127231"
/codon_start=1
/product="hypothetical protein LOC516026"
/protein_id="AAI05148.1"
/db_xref="GI:75773418"
/db_xref="GeneID:516026"
/translations="MSTVAPGCRSLDSRYRLCKKAAWGLVLEALAGVANASVA FMITLVLICKVDNSKRLPQFLGLVGLTFAFIITLDGGTPTFFLFG VLFALCFCLVHAFNLKLVGRQPLSMVLMLGLGFSLVQVDIIAIEYVLTWNRT NVNIFSELSPRRNEDFVMLLIYVLFMALFTLTASFQSGFTAWKRHGHYLTITI LSAIWMVATLLVPTLSPQWDDTILSSALVANGWVFLAVIAPFHLHTROQNPMD YPVEDAFQPEKQSYGVNRAYSQEGIIQGTTEGTSYLYAPSTYHFLQNRDSKD FSIPTVQTRVSPSYSDYEGNRDVS"

ORIGIN

Query Match 54.3%; Score 22.8; DB 14; Length 2069;
Best Local Similarity 71.4%; Pred. No. 98;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCGGCCCTTACAAGACTATGAAGTAAAG 42
|||||
Db 1153 GTCCAGACTCGGGTTAGCCCTTACGTGACTACGAAGGAAG 1194

RESULT 49

S75726

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REMARK

FEATURES

source

CDS

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

RESULT 50

S75726

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

FEATURES

source

CDS

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

RESULT 50

S75726

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

FEATURES

source

CDS

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

RESULT 50

S75726

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

FEATURES

source

CDS

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

S75726 Japanese encephalitis virus polyprotein gene, partial cds.
S75726 Japanese encephalitis virus polyprotein gene, partial cds.
S75726.1 GI:913060
Japanese encephalitis virus
Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
1 (bases 1 to 2379)
Hasegawa, H., Yoshida, M., Fujita, S. and Kobayashi, Y.
Comparison of structural proteins among antigenically different
Japanese encephalitis virus strains
Vaccine 12 (9), 841-844 (1994)
7975862
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 161325] from the original journal article.
Location/Qualifiers
1..2379
/organism="Japanese encephalitis virus"
/mol_type="genomic RNA"
/db_xref="taxon:11072"
/note="Nakayama-RFVL"
<1..>2379
/note="conceptual translation presented here differs from
translation in publication"
/codon_start=1
/product="polyprotein"
/protein_id="AA832599.2"
/db_xref="GI:9864807"
/translation="TKKPGPGKRAINMLKRLPRVFLVGVKRVMSLLDGRGPVR
FLVAFITVFKTALGRKAVKSVAMKHLTSFKRELGTLDIVVNRKGRKON
KRGNGESIMLAVIACAGAMKLSNFQGLKLTNNVDIADIVIVPNSKGRNC
WYRAIDVGYMCEDTITYECPKLTMGNDPEDVDCWCDNQVYVQYGRCTTRHSKRSR
SVVQTHGESSLVNKEAWLDSKATRYLMKTNWIVNPGVAFALAILGWLGSNG
QVVFITLLLVAPAYSFNCLGNGRDFIEGASGTWDLVLEGDSCLTIMANDKPTL
DYRMINIEAVQAEVSYCHASVTDISTVARCPTTGEAHNEKRDSSVYCKQGFDR
GWNGCGFLFGKSIDTKFCSTKAIGIQPENIKYEVGIFVHGTTTSENHNTSA
OVGASOAKFTVTPNAPSITLKLGYDEVTLDEPRSGLTAFVMTVGSKSFLVHR
EFNFDLALPWTPTSTANRELLMEFEAHATKOSVVALGSOEGLHQAALCAIIVE
YSSVKLTSLKCRKMDKALKKTGYCMCTEKFSFAPKNADTGHGTVIELSYSGS
DGCKIPIVSVASLNDMTPVGRVTVNPFVATSSANKVLVMEPEPFDSIYVVRGD
KQINQHHKAGSTLKAFSTTLKGAQLAALGDTAWDFSGIGVFNSIGKAVHQVFGG
APRTFLGGMSWITQGLMGALLMGMVGNARDRSIALAFLATGRVLFVFLATNVHA"

Query Match 54.3%; Score 22.8; DB 10; Length 2379;
Best Local Similarity 79.4%; Pred. No. 1e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CACGCTTGGCGGCCCTTACAAGACTATGAAG 37
|||||
Db 2082 CACGCTGGCGCAAGCCCTTTCAACGACTTTGAAG 2115
Japanese encephalitis virus Saigon polyprotein gene, partial cds.
JEU03696
U03696
U03696.1 GI:517409
Japanese encephalitis virus
Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Flavivirus; Japanese encephalitis virus group.
1 (bases 1 to 2436)
Ni, H. and Barrett, A. D.
Nucleotide and deduced amino acid sequence of the structural

protein genes of Japanese encephalitis viruses from different geographical locations
J. Gen. Virol. 76 (Pt 2), 401-407 (1995)
7844559
REFERENCE 2 (bases 1 to 2436)
Ni, H.
Direct Submission
Submitted (23-NOV-1993) Haolin Ni, Pathology, University of Texas Medical Branch, P-05, Galveston, TX 77555-0066, USA
JOURNAL
MEDICAL BRANCH, P-05, GALVESTON, TX 77555-0066, USA
FEATURES
source
1..2436
/organism="Japanese encephalitis virus"
/mol_type="genomic RNA"
/strain="Saigon"
/db_xref="taxon:11072"
/clone="pFRL9"
misc_feature
1..2436
/note="this sequence corresponds to nucleotides 28-2463 in the full JE virus genome"
CDS
69..>2436
/note="encodes C, prM, M, and E proteins"
/codon_start=1
/product="polyprotein"
/protein_id="AA67165.1"
/db_xref="GI:517410"
/translation="MTKPGGPKRINMLKRGLPVRVPLVGVKRVVMSLLDGRGPV RFLALITFFKTAALAPKALGRWKAKEVSAMKHLTSFKRELGLTLDIVANKRGKQ NKRGNESIMWLASLAVIACAGAKLSNFQGLKLTNNNTDADVIIPNPKGENRCWRAIDVGMCEDT IYECPLTNGNDPEDVDCWNOEVYQGRCTRHSKRSRSVSQTHGESLNV KKEAWLDSKATRYLMTKNTWIRNPGYALAAILGMLGNGNGRRWYFTLLILVA PAYFNCILGMGNRDFIEGASGATWDLVLEGDSCLTIMANDKPTLDVMINTEAVQLA EVRSYCVHASVTDISTVARCPTTGEAHNEKRADSSVCKQGFTRDGMNGCQLFGKGS IDTCAKFSCTSKAIGRTIQPENIKYEVGIFVHGTTTSENHNGYSAQVAGSAQAKTIVT FNAPSITLKLGDYGEVTLDCPRSGLNTAEFYMTVGSKFLVHREWFHDLALPPTPP SSTAWNRRELMFEFEAHATKQSVVVALGSQEGGLHQALAGAIIVSEYSSVKLTSGHLK CRLKMDKLALGTTTYGMCCTEKFSAKNPADTGHGTVTIELSYSGSDGCKPIPIVAS LNDMTPVGLRVTVPNFVATSSANSKVLVEMEPFGDSYIVVGRGDKQINHHMKAGST LGKAFSTTLKGAORLAALGDTAWDFSGISGVFNSIGKAVHQVFGCAFRTLFGEMCSWIT OGLGALLLWGVNARDRSIALAFLATGGVLVELATNVHADTGCAIDITRKEMRCSG IFVNDVPAWDRYKILPETPRSLAKIVHKAHKEGVCGRSVTRLEHQWNEAVRDELN VLLKENAVDLSVNVNPKVGRYSRSLMSTQEFKEMGWKAGKILIFAPELANSTFV VDGTEPCDPDEHRAWNSIEIDPFQGITSTRVMLKIREESTDECDGAIIGTAVKGHV AVHSLSYIESRYNDTWKLERAVEGKSCVTPETHLWDGDSVESLIIIPHTAGP KSKHNRGGYTKQNGPMDENGLVLDFOYCPGTVTITEDCGKGPSVRTTIDSGKLI TDWCRCSSLPBTPERTENGCMYGMERIPVRHDETTLVRSOVDAENGEMVDPFQGLL WFLQATVLRKRWARTLTIIPAVLGALLVLMGGITYTDLARYVYVVAFAEASGG DVHLHALIAVFKIQAPFALVMNMLSTRNTQENVVLGAFFHLASVDLQIGVHGLN AAAATKMMIVRAITPTTSSVTPVLLALTPGMRALYLDTRYIILLIIVIGICSLQBRKK TMAKKGAVLLGLALTSTGFSPTTAAAGLVCPNKKRGPATFELSAVGLMFAIVG GAELDIEMSISIPMLAGLMAVSYYVSGKATDMLERAADISWENDAAITGSSRLDV KLDDGDPHLIDDPCVPKWKVLMRSCIGLAALTIPWAIIPAAFGYWLTLKTTKRGGVF WDTPT"
ORIGIN
Query Match 54.3%; Score 22.8; DB 10; Length 2436;
Best Local Similarity 79.4%; Pred. No. 1e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
Db 2153 CACGCTGGCAAGCCCTTTCAACGACTTTGAAG 2186
RESULT 51
JEPVPT
LOCUS
DEFINITION
Japanese encephalitis virus capsid protein (3' end), membrane protein, envelope protein, nonstructural proteins NS1 and NS2, complete cds and nonstructural protein NS3 (3' end) mRNA.
ACCESSION
M16574
VERSION
M16574.1
KEYWORDS
4320 bp ss-RNA linear VRL 02-AUG-1993
SOURCE
ORGANISM
Japanese encephalitis virus
Japanese encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Flavivirus; Japanese encephalitis virus group.
REFERENCE
1 (bases 1 to 4320)
McAda, P.C., Mason, P.W., Schmaljohn, C.S., Dalrymple, J.M., Mason, T.L. and Fournier, M.J.
Partial nucleotide sequence of the Japanese encephalitis virus genome
JOURNAL
Virology 158 (2), 348-360 (1987)
PUBMED
3035787
COMMENT
Original source text: Japanese encephalitis virus (strain Nakaya), CDNA to viral RNA, passed 54 times in suckling mice.

Draft entry and computer-readable sequence for [1] kindly provided by D.Taylor, 29-SEP-1987.
Location/Qualifiers
1..4320
/organism="Japanese encephalitis virus"
/mol_type="genomic RNA"
/db_xref="taxon:11072"
<1..>4320
/note="polyprotein"
/codon_start=1
/protein_id="AAA46251.1"
/db_xref="GI:331337"
/translation="SVAMKHLTSPKRELGLTLDIVANKRKNKRGNGESIMWLASL AVIACAGAKLSNFPQGLKLTNNNTDADVIIPNPKGENRCWRAIDVGMCEDT IYECPLTNGNDPEDVDCWNOEVYQGRCTRHSKRSRSVSQTHGESLNV KKEAWLDSKATRYLMTKNTWIRNPGYALAAILGMLGNGNGRRWYFTLLILVA PAYFNCILGMGNRDFIEGASGATWDLVLEGDSCLTIMANDKPTLDVMINTEAVQLA EVRSYCVHASVTDISTVARCPTTGEAHNEKRADSSVCKQGFTRDGMNGCQLFGKGS IDTCAKFSCTSKAIGRTIQPENIKYEVGIFVHGTTTSENHNGYSAQVAGSAQAKTIVT FNAPSITLKLGDYGEVTLDCPRSGLNTAEFYMTVGSKFLVHREWFHDLALPPTPP SSTAWNRRELMFEFEAHATKQSVVVALGSQEGGLHQALAGAIIVSEYSSVKLTSGHLK CRLKMDKLALGTTTYGMCCTEKFSAKNPADTGHGTVTIELSYSGSDGCKPIPIVAS LNDMTPVGLRVTVPNFVATSSANSKVLVEMEPFGDSYIVVGRGDKQINHHMKAGST LGKAFSTTLKGAORLAALGDTAWDFSGISGVFNSIGKAVHQVFGCAFRTLFGEMCSWIT OGLGALLLWGVNARDRSIALAFLATGGVLVELATNVHADTGCAIDITRKEMRCSG IFVNDVPAWDRYKILPETPRSLAKIVHKAHKEGVCGRSVTRLEHQWNEAVRDELN VLLKENAVDLSVNVNPKVGRYSRSLMSTQEFKEMGWKAGKILIFAPELANSTFV VDGTEPCDPDEHRAWNSIEIDPFQGITSTRVMLKIREESTDECDGAIIGTAVKGHV AVHSLSYIESRYNDTWKLERAVEGKSCVTPETHLWDGDSVESLIIIPHTAGP KSKHNRGGYTKQNGPMDENGLVLDFOYCPGTVTITEDCGKGPSVRTTIDSGKLI TDWCRCSSLPBTPERTENGCMYGMERIPVRHDETTLVRSOVDAENGEMVDPFQGLL WFLQATVLRKRWARTLTIIPAVLGALLVLMGGITYTDLARYVYVVAFAEASGG DVHLHALIAVFKIQAPFALVMNMLSTRNTQENVVLGAFFHLASVDLQIGVHGLN AAAATKMMIVRAITPTTSSVTPVLLALTPGMRALYLDTRYIILLIIVIGICSLQBRKK TMAKKGAVLLGLALTSTGFSPTTAAAGLVCPNKKRGPATFELSAVGLMFAIVG GAELDIEMSISIPMLAGLMAVSYYVSGKATDMLERAADISWENDAAITGSSRLDV KLDDGDPHLIDDPCVPKWKVLMRSCIGLAALTIPWAIIPAAFGYWLTLKTTKRGGVF WDTPT"
mat_peptide
<1..159
/product="capsid protein"
sig_peptide
160..438
/note="membrane protein signal peptide"
mat_peptide
439..666
/product="membrane protein"
mat_peptide
667..2166
/product="envelope protein"
mat_peptide
2167..3408
/product="nonstructural protein NS1"
mat_peptide
3409..3903
/product="nonstructural protein ns2a"
mat_peptide
3904..4296
/product="nonstructural protein ns2b"
mat_peptide
4297..>4320
/product="nonstructural protein NS3"
ORIGIN
1908 bp upstream of SacI site.
Query Match 54.3%; Score 22.8; DB 10; Length 4320;
Best Local Similarity 79.4%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
Db 1869 CACGCTGGCAAGCCCTTTCAACGACTTTGAAG 1902
RESULT 52
AR003118
LOCUS
DEFINITION
Sequence 52 from patent US 5744140.
ACCESSION
AR003118
VERSION
AR003118.1
KEYWORDS
GI:3964377
SOURCE
Unknown.

[illegible]

ORIGIN

Query Match 54.3% Score 22.8; DB 10; Length 4512;

```
Best Local Similarity 79.4%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 7;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
||||| ||||| ||||| ||||| ||||| |||||
Db 2085 CACGCTGGCAAGGCTTTTCAACGACTTTGAAG 2118

RESULT 55
I08692 108692 4992 bp DNA linear PAT 02-DEC-1994
LOCUS Sequence 2 from Patent WO 8803032.
DEFINITION I08692
ACCESSION I08692.1 GI:588599
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4992)
McAda, P.C. and Mason, P.W.
DIAGNOSIS OF AND VACCINE FOR JAPANESE ENCEPHALITIS VIRUS AND
RELATED VIRUSES
JOURNAL Patent: WO 8803032-A 2 05-MAY-1988;
FEATURES
source
Location/Qualifiers
1..4992
/mol_type="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 54.3%; Score 22.8; DB 2; Length 4992;
Best Local Similarity 79.4%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
||||| ||||| ||||| ||||| ||||| |||||
Db 1869 CACGCTGGCAAGGCTTTTCAACGACTTTGAAG 1902

RESULT 56
BD064747 6536 bp DNA linear PAT 27-AUG-2002
LOCUS Expression vector containing cDNA encoding Japanese encephalitis
DEFINITION virus and vaccine using it.
ACCESSION BD064747.1 GI:22610350
VERSION
KEYWORDS JP 2001299355-A/3.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 6536)
Mason, P.W., Konishi, E., Yamaoka, M. and Kurane, I.
Expression vector containing cDNA encoding Japanese encephalitis
virus and vaccine using it
Patent: JP 2001299355-A 3 30-OCT-2001;
JOURNAL DIRECTOR GENERAL OF NATIONAL INSTITUTE OF INFECTIONS DISEASES,
ICHIRO KURANE, EIJI KONISHI
OS Artificial Sequence
PN JP 2001299355-A/3
PD 30-OCT-2001
PF 27-APR-2000 JP 2000126989
PT PETER W MASON, EIJI KONISHI, MASAKI YAMAOKA, ICHIRO KURANE PC
C12N15/09, A61K39/12, A61P31/12, C07K14/18, C12P21/02, PC
C12N15/09, C12R1.921, C12N15/00, C12N15/00, C12R1.92) CC Expression
vector containing cDNA encoding Japanese CC
encephalitis virus and
CC vaccine using it
FH Key Location/Qualifiers
FT source 1..6536
/mol_type="synthetic construct"
/mol_type="genomic DNA"
ORIGIN

Query Match 54.3%; Score 22.8; DB 2; Length 4992;
Best Local Similarity 79.4%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
||||| ||||| ||||| ||||| ||||| |||||
Db 2710 CACGCTGGCAAGGCTTTTCAACGACTTTGAAG 2743

RESULT 57
BD064684 7486 bp DNA linear PAT 27-AUG-2002
LOCUS Preparation of Flavivirus-like grains that are stable and non-toxic
DEFINITION as vaccine and diagnostic antigen.
ACCESSION BD064684
VERSION BD064684.1 GI:22610287
KEYWORDS JP 2001299336-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 7486)
Konishi, E. and Mason, P.W.
Preparation of Flavivirus-like grains that are stable and non-toxic
as vaccine and diagnostic antigen
Patent: JP 2001299336-A 1 30-OCT-2001;
JOURNAL EIJI KONISHI
OS Artificial Sequence
PN JP 2001299336-A/1
PD 30-OCT-2001
PF 28-APR-2000 JP 2000130250
PT EIJI KONISHI, PETER W MASON
PC C12N15/09, A61K39/12, A61P31/12, C07K14/18, C12N15/09, C12P21/02, CC
C12N15/00, C12R1.91, C12N15/00, C12N15/00, C12N15/00, C12R1.91) CC
Preparation of Flavivirus-like grains that are stable and non- CC
toxic as
CC vaccine and diagnostic antigen
FH Key Location/Qualifiers
FT source 1..7486
/mol_type="synthetic construct"
/mol_type="genomic DNA"
/mol_type="taxon:32630"
ORIGIN

Query Match 54.3%; Score 22.8; DB 2; Length 7486;
Best Local Similarity 79.4%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
||||| ||||| ||||| ||||| ||||| |||||
Db 2710 CACGCTGGCAAGGCTTTTCAACGACTTTGAAG 2743

RESULT 58
BD064685 7486 bp DNA linear PAT 27-AUG-2002
LOCUS Preparation of Flavivirus-like grains that are stable and non-toxic
DEFINITION as vaccine and diagnostic antigen.
ACCESSION BD064685
VERSION BD064685.1 GI:22610288
KEYWORDS JP 2001299336-A/2.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 7486)
Konishi, E. and Mason, P.W.
Preparation of Flavivirus-like grains that are stable and non-toxic
as vaccine and diagnostic antigen
/mol_type="genomic DNA"
ORIGIN
```



```

JOURNAL      Patent: JP 2001299336-A 2 30-OCT-2001;
COMMENT      EIJI KONISHI
              OS Artificial Sequence
              PN JP 2001299336-A/2
              PD 30-OCT-2001
              PF 28-APR-2000 JP 2000130250
              PI EIJI KONISHI, PETER W MASON
              PC C12N5/10.A61K39/12.A61P31/12.C07K14/18.C12N15/09.C12P21/02//
              PC (C12N5/10.C12R1.91).C12N5/00.C12N15/00.(C12N5/00.C12R1.91) CC
              Preparation of Flavivirus-like grains that are stable and non-CC
              toxic as
              CC vaccine and diagnostic antigen
              FH Key Location/Qualifiers
              FT source 1..7486
              FT /organism='Artificial Sequence'.
              FEATURES
                source
                  Location/Qualifiers
                    1..7486
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
              ORIGIN
                Query Match 54.3%; Score 22.8; DB 2; Length 7486;
                Best Local Similarity 79.4%; Pred. No. 1.5e+02;
                Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCGAGCCCTTACAAGACTATGAAG 37
   ||||| ||| ||||| ||| ||||| ||| |||||
Db 2710 CACGCTGGGCAAGCCCTTTCAACGACTTTGAAG 2743

RESULT 59
LOCUS      BD064746 7486 bp DNA linear PAT 27-AUG-2002
DEFINITION Expression vector containing cDNA encoding Japanese encephalitis
            virus and vaccine using it.
ACCESSION  BD064746
VERSION     BD064745.1 GI:22610349
KEYWORDS    JP 2001299355-A/2.
SOURCE      synthetic construct
            ORGANISM
              other sequences; artificial sequences.
REFERENCE   1 (bases 1 to 7486)
            Mason, P.W., Konishi, E., Yamaka, M. and Kurane, I.
            Expression vector containing cDNA encoding Japanese encephalitis
            virus and vaccine using it
            Patent: JP 2001299355-A 2 30-OCT-2001;
            DIRECTOR GENERAL OF NATIONAL INSTITUTE OF INFECTIONS DISEASES,
            ICHIRO KURANE, EIJI KONISHI
            OS Artificial Sequence
            PN JP 2001299355-A/2
            PD 30-OCT-2001
            PF 27-APR-2000 JP 2000126989
            PI PETER W MASON, EIJI KONISHI, MASAKI YAMAOKA, ICHIRO KURANE PC
            C12N15/09.A61K39/12.A61P31/12.A61P31/12./C07K14/18.C12P21/02.PC
            (C12N15/09.C12R1.92).C12N15/00.(C12N15/00.C12R1.92) CC Expression
            vector containing cDNA encoding Japanese CC
            encephalitis virus and
            CC vaccine using it
            FH Key Location/Qualifiers
            FT source 1..7486
            FT /organism='Artificial Sequence'.
            FEATURES
              source
                Location/Qualifiers
                  1..7486
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
              ORIGIN
                Query Match 54.3%; Score 22.8; DB 2; Length 7486;
                Best Local Similarity 79.4%; Pred. No. 1.5e+02;
                Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

JOURNAL      Patent: JP 2001299336-A 2 30-OCT-2001;
COMMENT      EIJI KONISHI
              OS Artificial Sequence
              PN JP 2001299336-A/2
              PD 30-OCT-2001
              PF 28-APR-2000 JP 2000130250
              PI EIJI KONISHI, PETER W MASON
              PC C12N5/10.A61K39/12.A61P31/12.C07K14/18.C12N15/09.C12P21/02//
              PC (C12N5/10.C12R1.91).C12N5/00.C12N15/00.(C12N5/00.C12R1.91) CC
              Preparation of Flavivirus-like grains that are stable and non-CC
              toxic as
              CC vaccine and diagnostic antigen
              FH Key Location/Qualifiers
              FT source 1..7486
              FT /organism='Artificial Sequence'.
              FEATURES
                source
                  Location/Qualifiers
                    1..7486
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
              ORIGIN
                Query Match 54.3%; Score 22.8; DB 2; Length 7486;
                Best Local Similarity 79.4%; Pred. No. 1.5e+02;
                Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCGAGCCCTTACAAGACTATGAAG 37
   ||||| ||| ||||| ||| ||||| ||| |||||
Db 2710 CACGCTGGGCAAGCCCTTTCAACGACTTTGAAG 2743

RESULT 60
LOCUS      JEVBEICG 10976 bp ss-RNA linear VRL 21-NOV-1995
DEFINITION Japanese encephalitis virus (strain Beijing-1), complete genome.
ACCESSION  L48961
VERSION     L48961.1 GI:1066797
KEYWORDS    .
SOURCE      Japanese encephalitis virus
            ORGANISM
              Japanese encephalitis virus
              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Flavivirus; Japanese encephalitis virus group.
              1 (sites)
              Hashimoto, H., Nomoto, A., Watanabe, K., Mori, T., Takezawa, T.,
              Aizawa, C., Takegami, T. and Hiramatsu, K.
              Molecular cloning and complete nucleotide sequence of the genome of
              Japanese encephalitis virus Beijing-1 strain
              Virus Genes 1 (3), 305-317 (1988)
              PUBMED 3245133
              COMMENT Original source text: Japanese encephalitis virus (strain
              Beijing-1) RNA.
              FEATURES
                source
                  Location/Qualifiers
                    1..10976
                    /organism="Japanese encephalitis virus"
                    /mol_type="genomic RNA"
                    /strain="Beijing-1"
                    /db_xref="taxon:11072"
              ORIGIN
                Query Match 54.3%; Score 22.8; DB 10; Length 10976;
                Best Local Similarity 79.4%; Pred. No. 1.8e+02;
                Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCGAGCCCTTACAAGACTATGAAG 37
   ||||| ||| ||||| ||| ||||| ||| |||||
Db 2180 CACGCTGGGCAAGCCCTTTCAACGACTTTGAAG 2213

RESULT 61
LOCUS      AC025723/c 77422 bp DNA linear INV 07-SEP-2005
DEFINITION Caenorhabditis elegans cosmid Y54F10A1, complete sequence.
ACCESSION  AC025723
VERSION     AC025723.2 GI:13559726
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
            ORGANISM
              Caenorhabditis elegans
              Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
              Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
              1 (bases 1 to 77422)
              C. elegans Sequencing Consortium
              Genome sequence of the nematode C. elegans: a platform for
              investigating biology
              Science 282 (5396), 2012-2018 (1998)
              PUBMED 9851916
              REFERENCE
                2 (bases 1 to 77422)
                Ryan, E., Courtney, L. and Yoakum, M.
                The sequence of C. elegans cosmid Y54F10A1
                Unpublished (2001)
              REFERENCE
                3 (bases 1 to 77422)
                Waterston, R.H.
                Direct Submission
                Submitted (12-MAR-2000) Genome Sequencing Center, Washington
                University School of Medicine, 4444 Forest Park Parkway, St. Louis,
                MO 63108, USA
              REFERENCE
                4 (bases 1 to 77422)
                Waterston, R.
                Direct Submission

```



```

YK810D07.3, YK810D07.5, YK850B07.3, YK850B07.5,
YK879D06.3, YK879D06.5, YK880C07.3, YK880C07.5,
YK896A02.3, YK896A02.5, YK900C05.3, YK900C05.5,
YK1005E12.3, YK1005E12.5, YK1012G08.5, YK1015F05.5,
YK1024E05.3, YK1024E05.5, YK1051B01.3, YK1051B01.5,
YK1054E06.3, YK1054E06.5, YK1059B11.3, YK1059B11.5,
YK1071D09.3, YK1071D09.5, YK1073A03.3, YK1073A03.5,
YK1076B05.3, YK1076B05.5, YK1090C03.3, YK1090C03.5,
YK1101C09.3, YK1101C09.5, YK1124H09.3, YK1124H09.5,
YK1135B09.3, YK1135B09.5, YK1139A10.3, YK1139A10.5,
YK1186H04.3, YK1186H04.5, YK1189G02.3, YK1189G02.5,
YK1198C07.3, YK1198C07.5, YK1202C05.3, YK1202C05.5,
YK1207G10.3, YK1207G10.5, YK1219C01.5,
/codon_start=1
/product="Hypothetical protein Y54F10AM.8"
/protein_id="AK29943.1"
/db_xref="GI:13559738"
/db_xref="WormBase:Y54F10AM.8"
/db_xref="WormBase:WBGene00021852"
/tranlation="MKLLFLFLGLIFAVQEKPLDNNRPVPEQLNDHSSAKFDYTY
VSVUNSTDETLDDIVYAECKNAASRVALKYKSNQVNTGNGILEIEFASHSDVQ
AGVAGBELRLQIYHYRNTIETMCNNHTLFCRLYIYIQQLDMMRSQVQANPP
TDFVRQVNLTAQITGIDAYSKRNLTPETIGFLPIYMQLAGMDLNLKNTA
DMPEYEGRCRSGFKLAPCNKMFMAHVSMSLSMMQSVLKIYKFGYDNEVPGHIV
TFSGYPGVLISDDVITTSAGLTSIETITIAFNQTLTYDKFMKPEQVHCWIRMSLN
54.3%; Score 22.8; DB 13; Length 77422;
Query Match 71.4%; Pred. No. 3.5e+02;
Best Local Similarity 0; Mismatches 12; Indels 0; Gaps 0;
Matches 30; Conservative 0;

QY 1 GCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAG 42
Db 51132 GCAGATGCTTGTTCGAGCCATTGGAAGACGCTGAAGGATAG 51091

RESULT 62
AC121911/c
LOCUS AC121911.3 GI:27452957
DEFINITION Mus musculus BAC clone RP24-181L3 from chromosome 1, complete
sequence.
AC121911
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 141695)
Scott, K., Kalicki, J., Dignan, G., Spalding, L., Veizer, J. and
Mangiapanello, J.
The sequence of Mus musculus BAC clone RP24-181L3
Unpublished (2001)
2 (bases 1 to 141695)
Wilson, R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 141695)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 141695)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (01-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 141695)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (01-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 141695)

```

AUTHORS TITLE JOURNAL COMMENT

Wilson, R.
Direct Submission
Submitted (11-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 1, 2003 this sequence version replaced gi:22475765.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0181L03

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES source

```

1. 141695
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="1"
/map="1"
/clone="RP24-181L3"
/clone_lib="RPCI-24"
2189..2252
/rpt_family="L1"
2368..2493
/rpt_family="Alu"
3981..4084
/rpt_family="Alu"
4245..4304
/rpt_family="ID"
4361..4444
/rpt_family="Alu"
4454..4690
/rpt_family="MER1_type"
4857..4958
/rpt_family="ACHobo"
6389..6709
/rpt_family="L1"
6721..6816
/rpt_family="Alu"
7227..7275
/rpt_family="Alu"
7276..7413
/rpt_family="Alu"
7414..7439

```

```

/rpt_family="Alu"
repeat_region      8409..8486 /rpt_family="B4"
repeat_region      28033..28256 /rpt_family="L1"
repeat_region      9018..9204 /rpt_family="B2"
repeat_region      28435..28497 /rpt_family="B2"
repeat_region      9638..9745 /rpt_family="L1"
repeat_region      29406..29695 /rpt_family="ERVK"
repeat_region      29685..29762 /rpt_family="ERVK"
repeat_region      29763..30127 /rpt_family="MaLR"
repeat_region      30092..30280 /rpt_family="ERVK"
repeat_region      30921..30977 /rpt_family="B4"
repeat_region      30978..31088 /rpt_family="Alu"

Query Match          54.3%; Score 22.8; DB 6; Length 141695;
Best Local Similarity 71.4%; Pred.No. 4.3e+02;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy   1 GCCACGCTTGCGGGAGCCCTTACAAGAAGACTATGAAGTAAG 42
Db   136640 GCCCAGCATGGCTGCCTTTAAAAAATAAAAAAAAAAAG 136599

RESULT 63
AC142410/c
LOCUS       AC142410                145213 bp    DNA     linear   ROD 28-JAN-2005
DEFINITION Mus musculus BAC clone RP24-173I18 from 1, complete sequence.
ACCESSION   AC142410
VERSION     AC142410.4 GI:51854709
KEYWORDS
SOURCE      Mus musculus
            Mus musculus (house mouse)
ORGANISM    Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1 (bases 1 to 145213)
AUTHORS    Levy,A., Bleilick,I.D. and Meyer,R.
TITLE      The sequence of Mus musculus BAC clone RP24-173I18
JOURNAL    Unpublished (2001)
REFERENCE   2 (bases 1 to 145213)
AUTHORS    McPherson,J.D. and Waterston,R.H.
TITLE      Direct Submission
JOURNAL    Submitted (28-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3 (bases 1 to 145213)
AUTHORS    Wilson,R.K.
TITLE      Direct Submission
JOURNAL    Submitted (13-JUL-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   4 (bases 1 to 145213)
AUTHORS    Wilson,R.K.
TITLE      Direct Submission
JOURNAL    Submitted (02-SEP-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   5 (bases 1 to 145213)
AUTHORS    Wilson,R.K.
TITLE      Direct Submission
JOURNAL    Submitted (28-JAN-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     On Sep 2, 2004 this sequence version replaced gi:50253691.
           ----- Genome Center
           Center: Washington University Genome Sequencing Center
           Web site: http://genome.wustl.edu
           Contact: submissions@wustl.edu
           Summary Statistics
           Center project name: M BB0173118

```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source	Location/Qualifiers
1. .145213	/organism="Mus musculus"
	/mol_type="Genomic DNA"
	/db_xref="taxon:10090"
	/chromosomes="1"
	/map="1"
	/clone="RP24-173118"
	/clone_lib="RPCI-24"
4. .2142	/rpt_family="L1"
7798..8187	/rpt_family="ERVK"
8188..8662	/rpt_family="L1"
8667..10718	/rpt_family="L1"
12405..12534	/rpt_family="Alu"
13132..13246	/rpt_family="Alu"
13441..13507	/rpt_family="ID"
13742..13883	/rpt_family="Alu"
13948..14263	/rpt_family="MaLR"
14264..14437	/rpt_family="ERVL"
14576..14702	/rpt_family="Alu"
15146..15952	/rpt_family="L1"
15953..16137	/rpt_family="B2"
16138..16688	/rpt_family="L1"
16686..19681	/rpt_family="L1"
21390..21519	/rpt_family="Alu"
22118..22210	/rpt_family="Alu"
22418..22488	/rpt_family="ID"
22683..22824	/rpt_family="Alu"
repeat_region	22891..23382 /rpt_family="MaLR"
repeat_region	23497..23632 /rpt_family="Alu"
repeat_region	24137..24255 /rpt_family="Alu"
repeat_region	24275..24403 /rpt_family="Alu"
repeat_region	24412..24662 /rpt_family="L1"
repeat_region	24698..24807 /rpt_family="B4"
repeat_region	25032..25187 /rpt_family="L1"
repeat_region	25192..25633 /rpt_family="L1"
repeat_region	25751..26167 /rpt_family="ERVK"
repeat_region	26168..26350 /rpt_family="B2"
repeat_region	26351..26480 /rpt_family="ERVK"
repeat_region	26508..26655 /rpt_family="Alu"
repeat_region	26656..26999 /rpt_family="ERVK"
repeat_region	27000..27315 /rpt_family="MaLR"
repeat_region	27531..27934 /rpt_family="ERVK"
repeat_region	28312..28365 /rpt_family="MaLR"
repeat_region	28372..28563 /rpt_family="B2"
repeat_region	28564..28705 /rpt_family="B4"
repeat_region	29337..29460 /rpt_family="B4"
repeat_region	29506..30142 /rpt_family="L1"
repeat_region	30245..30307 /rpt_family="4.5SRNA"
repeat_region	30260..30327 /rpt_family="Alu"
repeat_region	31197..31340 /rpt_family="Alu"
repeat_region	33345..33443 /rpt_family="Alu"
repeat_region	34066..34180 /rpt_family="Alu"
repeat_region	34374..34429 /rpt_family="ID"
repeat_region	34670..34811 /rpt_family="Alu"
repeat_region	34876..35267 /rpt_family="MaLR"
repeat_region	35268..35365 /rpt_family="ERVL"
repeat_region	35483..35627 /rpt_family="Alu"
repeat_region	36082..36881 /rpt_family="L1"
repeat_region	36882..37067 /rpt_family="B2"
repeat_region	37068..37618 /rpt_family="L1"
repeat_region	37616..38978 /rpt_family="L1"
repeat_region	38982..39335 /rpt_family="ERVK"
repeat_region	39392..39574 /rpt_family="ERVK"
repeat_region	39575..40131

```

/rpt_family="ERVK"
40132..40409
/rpt_family="ERVK"
40409..41158
/rpt_family="ERVK"
41161..41587
/rpt_family="ERVK"
41588..43274
/rpt_family="L1"
43660..44187
/rpt_family="L1"

Query Match      54.3%; Score 22.8; DB 6; Length 145213;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 28332 GCCACAGTGGCTGAGCCTTTAAAAAAMAAAAAAMAAAAAG 28291

RESULT 64
AC095186/c
LOCUS          247690 bp      DNA      linear      HTG 13-NOV-2002
DEFINITION    Rattus norvegicus clone CH230-9B21, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION     AC095186
VERSION       AC095186.6 GI:24941004
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 247690)
Muzny,D,Marie, Metsker,M,Lee, Abranzone,S, Adams,C, Alder,J,,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,,
Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,,
Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A,,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,,
Gebregregis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,,
Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,,
Hollins,B, Howells,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A,,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jollivet,A,,
Karkathi,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,,
Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,,
Lorensuheva,L, Loulseghe,H, Lozado,R,J, Lu,X, Ma,J,,
Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,,
Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,,
Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,,
Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L,,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,, Parks,K,,
Nwaokemele,O, Okwuonu,G, Olarnpunsagoon,A, Pal,S,, Parks,K,,
Pacteriak,S, Paul,H, Perez,A, Perez,L, Pfanckoch,C,,
Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L-L,,
Fuazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,,
Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,,
Sanders,W, Savary,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,,
Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D,,
Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Soosa,J,,
Stearns,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C,,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umani,K,,
Valas,R, Vera,V, Villalana,D, Waldron,L, Walker,B, Wang,J,,
Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,,
Williams,G, Willson,R, Wlarczyk,R, Wooden,H, Worley,K,,
Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,,
Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von
Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,,
Weinstock,G, and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 247690)
Worley,K.C.
Direct Submission
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 247690)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:22772986.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDBT
Center clone name: CH230-9B21
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 215114 bases at least Q40
Consensus quality: 217886 bases at least Q30
Consensus quality: 219758 bases at least Q20
Estimated insert size: 221396; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 101341: contig of 101341 bp in length
* 101342 101441: gap of unknown length
* 101442 219685: contig of 118244 bp in length
* 219686 219785: gap of unknown length
* 219786 247690: contig of 27905 bp in length.
Location/Qualifiers
1..247690
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

FEATURES
source

```



```

/db xref="taxon:10116"
/clone="CH230-9B21"
misc_feature
1..1604
/notes="wgs contig"
gap
101342..101441
/estimated_length=unknown
219686..219785
/estimated_length=unknown

ORIGIN
Query Match      54.3%; Score 22.8; DB 12; Length 247690;
Best Local Similarity 79.4%; Pred. No. 5.2e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 200484 TAGGACCAACACTTTCAAGAGCTTTGAAGTAAAG 200451

RESULT 65
AC137204
LOCUS
DEFINITION
Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
***, 4 unordered pieces.
ACCESSION
AC137204.1 GI:25073092
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 249669)

REFERENCE
AUTHORS
Muzny, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleaveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

```

Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tincey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 249669)
Rat Genome Sequencing Consortium.
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: K2PB
Center clone name: CH230-unknown
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 20161 bases at least Q40
Consensus quality: 206093 bases at least Q30
Consensus quality: 208911 bases at least Q20
Estimated insert size: 209892; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 103590: contig of 103590 bp in length
* 103591 103890: gap of unknown length
* 103691 218904: contig of 115214 bp in length
* 218905 219004: gap of unknown length
* 219005 248412: contig of 29408 bp in length
* 248413 248512: gap of unknown length
* 248513 249669: contig of 1157 bp in length.
Location/Qualifiers
1..249669
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-unknown"
1..1761
/note="wgs contig"
4671..6761
/note="wgs contig"
103591..103690

FEATURES
source
misc_feature
misc_feature
gap

```

/estimated_length=unknown
218905. .219004
/estimated_length=unknown
248413. .248512
/estimated_length=unknown

ORIGIN
Query Match          54.3%; Score 22.8; DB 12; Length 249669;
Best Local Similarity 79.4%; Pred. No. 5.2e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTGGCGGACCCCTTACAAGACTATGACGTAAG 42
Db 123020 TAGGACCAACACTTTCAAAGACTTTGAAGTAAAG 123053

RESULT 66
AC006875      298216 bp      DNA      linear      HTG 24-FEB-1999
Caenorhabditis elegans clone Y54F10x, *** SEQUENCING IN PROGRESS
***, 25 unordered pieces.
AC006875
AC006875.1 GI:4263473
HTG; HTGS PHASE1
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 298216)
Waterston.R.H.
The sequence of Caenorhabditis elegans clone
Unpublished
2 (bases 1 to 298216)
Waterston.R.H.
Direct Submission
Submitted (24-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2000: contig of 2000 bp in length
* 2001 2015: gap of unknown length
* 2016 3995: contig of 1980 bp in length
* 3996 4010: gap of unknown length
* 4011 7604: contig of 3594 bp in length
* 7605 7619: gap of unknown length
* 7620 10290: contig of 2671 bp in length
* 10291 10305: gap of unknown length
* 10306 14793: contig of 4487 bp in length
* 14793 14807: gap of unknown length
* 14808 18212: contig of 3405 bp in length
* 18213 18227: gap of unknown length
* 18228 21270: contig of 3043 bp in length
* 21271 21285: gap of unknown length
* 21286 24063: contig of 2777 bp in length
* 24063 27088: gap of unknown length
* 27088 27102: gap of unknown length
* 27103 32663: contig of 5560 bp in length
* 32663 32677: gap of unknown length
* 32678 38717: gap of unknown length
* 38717 42014: contig of 3010 bp in length
* 42014 42028: gap of unknown length
* 42029 47981: contig of 5952 bp in length
* 47981 47995: gap of unknown length
* 47996 52552: contig of 4557 bp in length

/estimated_length=unknown
218905. .219004
/estimated_length=unknown
248413. .248512
/estimated_length=unknown

FEATURES
source
1. 298216
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/db_xref="taxon:6239"
/clone="Y54F10x"
2001. .2015
/estimated_length=unknown
3996. .4010
/estimated_length=unknown
7605. .7619
/estimated_length=unknown
10291. .10305
/estimated_length=unknown
14793. .14807
/estimated_length=unknown
18213. .18227
/estimated_length=unknown
21271. .21285
/estimated_length=unknown
24083. .24077
/estimated_length=unknown
27088. .27102
/estimated_length=unknown
32663. .32677
/estimated_length=unknown
38717. .38731
/estimated_length=unknown
42014. .42028
/estimated_length=unknown
47981. .47995
/estimated_length=unknown
52553. .52567
/estimated_length=unknown
59007. .59021
/estimated_length=unknown
66562. .66576
/estimated_length=unknown
73618. .73632
/estimated_length=unknown
84893. .84907
/estimated_length=unknown
97462. .97476
/estimated_length=unknown
110333. .110347
/estimated_length=unknown
128092. .128106
/estimated_length=unknown
166559. .166573
/estimated_length=unknown
201396. .201410
/estimated_length=unknown
52553 52567: gap of unknown length
59006 59021: contig of 6439 bp in length
59021 66561: contig of 7540 bp in length
66561 66576: gap of unknown length
66576 73617: contig of 7041 bp in length
73617 73632: gap of unknown length
84892 84907: contig of 11260 bp in length
84907 97461: contig of 12554 bp in length
97461 97476: gap of unknown length
110332 110347: contig of 12856 bp in length
110347 128091: contig of 17744 bp in length
128091 128106: gap of unknown length
128107 166558: contig of 38452 bp in length
166558 166573: gap of unknown length
201395 201410: gap of unknown length
201410 275281: contig of 73871 bp in length
275281 275296: gap of unknown length
275296 298216: contig of 22920 bp in length.
Location/Qualifiers
1. 298216
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/db_xref="taxon:6239"
/clone="Y54F10x"
2001. .2015
/estimated_length=unknown
3996. .4010
/estimated_length=unknown
7605. .7619
/estimated_length=unknown
10291. .10305
/estimated_length=unknown
14793. .14807
/estimated_length=unknown
18213. .18227
/estimated_length=unknown
21271. .21285
/estimated_length=unknown
24083. .24077
/estimated_length=unknown
27088. .27102
/estimated_length=unknown
32663. .32677
/estimated_length=unknown
38717. .38731
/estimated_length=unknown
42014. .42028
/estimated_length=unknown
47981. .47995
/estimated_length=unknown
52553. .52567
/estimated_length=unknown
59007. .59021
/estimated_length=unknown
66562. .66576
/estimated_length=unknown
73618. .73632
/estimated_length=unknown
84893. .84907
/estimated_length=unknown
97462. .97476
/estimated_length=unknown
110333. .110347
/estimated_length=unknown
128092. .128106
/estimated_length=unknown
166559. .166573
/estimated_length=unknown
201396. .201410
/estimated_length=unknown
```

```
/estimated_length=unknown
275282..275296
/estimated_length=unknown

Query Match      54.3%; Score 22.8; DB 12; Length 298216;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

ORIGIN
1 GCCACGCTGGCGGACCCCTTACAAGACTATGAAGTAAG 42
2 ||||| ||||| ||||| ||||| ||||| |||||
3 GCATGCTTGTTCGAGCCATTGGAAGAGCGCTGAAGGATAG 152602

RESULT 67
AC141041
LOCUS      Rattus norvegicus clone CH230-473J6, *** SEQUENCING IN PROGRESS
DEFINITION *** 61 unordered pieces.
ACCESSION AC141041
VERSION    1 GI:28875900
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE  1 (bases 1 to 161575)
            Muzny, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
            Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
            Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
            Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
            Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
            Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
            Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
            Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
            Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
            Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
            Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
            Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
            Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
            Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
            Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
            Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
            Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
            Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
            Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
            Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
            Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
            Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
            Kowig, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
            Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
            Lorensuhera, L., Louleghed, H., Lozado, R., Lu, X., Lu, X., Ma, J.,
            Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
            Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
            Manthey, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
            Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
            Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
            Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G.,
            Olarungasagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
            Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,
            Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,
            Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
            Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
            Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S.,
            Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
            Sisson, I., Sitter, C.D., Smajd, D., Speed, A., Sodergren, E.,
            Song, X.-Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
            Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
            Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D.,
            Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
            Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R.,
            Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
```

```
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R.,
Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 161575)
Worley, K.C.
Direct Submission
Submitted (07-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161575)
Worley, K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GXIU
Center clone name: CH230-473J6
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 130671 bases at least Q40
Consensus quality: 139673 bases at least Q30
Consensus quality: 145030 bases at least Q20
Estimated insert size: 13121; sum-of-contigs estimation
Quality coverage: 2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1276: contig of 1276 bp in length
* 1277 1376: gap of unknown length
* 1377 2670: contig of 1294 bp in length
* 2671 2770: gap of unknown length
* 2771 4049: contig of 1279 bp in length
* 4050 4149: gap of unknown length
* 4150 5234: contig of 1085 bp in length
* 5235 5334: gap of unknown length
* 5335 6376: contig of 1041 bp in length
* 6376 6475: gap of unknown length
* 6476 7950: contig of 1475 bp in length
* 7951 8050: gap of unknown length
* 8051 10029: contig of 1979 bp in length
* 10030 10129: gap of unknown length
* 10130 11297: contig of 1168 bp in length
* 11298 11397: gap of unknown length
* 11398 12571: contig of 1174 bp in length
* 12572 12671: gap of unknown length
* 12672 13872: contig of 1201 bp in length
* 13873 13972: gap of unknown length
* 13973 15359: contig of 1387 bp in length
* 15360 15459: gap of unknown length
* 15460 16729: contig of 1270 bp in length
* 16730 16829: gap of unknown length
* 16830 17925: contig of 1095 bp in length
* 17925 18025: gap of unknown length
* 18025 19369: contig of 1345 bp in length
* 19370 19469: gap of unknown length
```

```

* 19470 21151: contig of 1682 bp in length
* 21152 21151: gap of unknown length
* 21252 22627: contig of 1376 bp in length
* 22628 22727: gap of unknown length
* 22728 24187: contig of 1460 bp in length
* 24188 24287: gap of unknown length
* 24288 26783: contig of 2496 bp in length
* 26784 26883: gap of unknown length
* 26884 27222: contig of 1039 bp in length
* 27923 28022: gap of unknown length
* 28023 29392: contig of 1270 bp in length
* 29393 31312: gap of unknown length
* 31313 31412: contig of 1920 bp in length
* 31413 32119: contig of 1107 bp in length
* 32120 32619: gap of unknown length
* 32620 34673: contig of 2054 bp in length
* 34674 34773: gap of unknown length
* 34774 36319: contig of 1546 bp in length
* 36320 36419: gap of unknown length
* 36420 37971: contig of 1552 bp in length
* 37972 38071: gap of unknown length
* 38072 40564: contig of 2493 bp in length
* 40565 40664: gap of unknown length
* 40665 42116: contig of 1352 bp in length
* 42017 42116: gap of unknown length
* 42117 43834: contig of 1718 bp in length
* 43835 43934: gap of unknown length
* 43935 45295: contig of 1361 bp in length
* 45296 45395: gap of unknown length
* 43996 47677: contig of 2282 bp in length
* 47678 47777: gap of unknown length
* 47778 49906: contig of 2129 bp in length
* 49907 50006: gap of unknown length
* 50007 52338: contig of 2232 bp in length
* 52339 53000: gap of unknown length
* 53001 53900: gap of unknown length
* 53901 55788: contig of 1888 bp in length
* 55789 55889: gap of unknown length
* 55889 58900: contig of 2912 bp in length
* 58901 62150: contig of 3250 bp in length
* 62151 62250: gap of unknown length
* 62251 64203: contig of 1953 bp in length
* 64204 64303: gap of unknown length
* 64304 67105: contig of 2802 bp in length
* 67106 67205: gap of unknown length
* 67206 69322: contig of 2727 bp in length
* 69333 70032: gap of unknown length
* 70033 72191: contig of 2159 bp in length
* 72192 72291: gap of unknown length
* 72292 75506: contig of 3215 bp in length
* 75507 75607: gap of unknown length
* 75607 77496: contig of 1890 bp in length
* 77497 77596: gap of unknown length
* 77597 80560: contig of 2964 bp in length
* 80561 80660: gap of unknown length
* 80661 83785: contig of 3125 bp in length
* 83786 83885: gap of unknown length
* 83886 86387: contig of 2502 bp in length
* 86388 86487: gap of unknown length
* 86488 89888: contig of 3401 bp in length
* 89889 92744: gap of unknown length
* 92745 92844: gap of unknown length
* 92845 97427: contig of 4583 bp in length
* 97428 97527: gap of unknown length
* 97528 100290: contig of 2763 bp in length
* 100291 100390: gap of unknown length

```

Query Match 53.8%; Score 22.6; DB 12; Length 161575;
 Best Local Similarity 75.7%; Pred. No. 5.5e+02;

```

Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 5 ACCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 41
Db 21098 AGTTTGGCCGAGCCCTTTCAATTCTATAAAGGAA 21134

RESULT 68
AC166012/c
LOCUS
DEFINITION
Oryctolagus cuniculus clone LB1-138P6, WORKING DRAFT SEQUENCE, 5
ORDERED PIECES.
AC166012
AC166012.2 GI:71980158
HTG: HTGS_PHASE2; HTGS_DRAFT.
Oryctolagus cuniculus (rabbit)
SOURCE
ORGANISM
Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha;
Leporidae; Oryctolagus.
1 (bases 1 to 173071)
Antonellis,A., Ayele,K., Bass,D., Benjamin,B., Bera,J.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G.,
Coleman,H., Engle,J., Franks,S., Fuksenko,T., Gestole,M.,
Greene,A., Guan,X., Gupta,J., Gurson,N., Haghighi,P., Han,J.,
Hansen,N., Ho,S.-L., Hu,P., Hunter,G., Hurie,B., Idol,J.R.,
Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q., Legaspi,R.,
Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C.,
Maskeri,B., McDowell,J., Mojidi,H.A., Mullikin,J.C., Park,M.,
Portnoy,M.E., Prasad,A., Puri,O., Rantiz,K., Reddix-Dugue,N.,
Sante,A., Schandler,K., Schueler,M.G., Sison,C., Stantripop,S.,
Taye,A., Thomas,J.W., Thomas,P.J., Teipouri,V., Ung,L., Vogt,J.L.,
Wetherby,K.D., Withers,T.R., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 173071)
Green,E.D.
Direct Submission
Submitted (20-JUL-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
3 (bases 1 to 173071)
Green,E.D.
Direct Submission
Submitted (09-AUG-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Aug 9, 2005 this sequence version replaced gi:71037466.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgrl.nih.gov
----- Project Information
Center project name: keo
Center clone name: 138P06

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is generally based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 172156 bases at least Q40
Consensus quality: 172470 bases at least Q30

```

Consensus quality: 172582 bases at least Q20
 Insert size: 195000; agarose-fp
 Insert size: 172671; sum-of-contigs
 Quality coverage: 9.00x in Q20 bases; agarose-fp
 Quality coverage: 10.16x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 2231: contig of 2231 bp in length
 * 2232 2331: gap of unknown length
 * 2232 6543: contig of 4212 bp in length
 * 6544 6643: gap of unknown length
 * 6544 94753: contig of 88110 bp in length
 * 94754 94853: gap of unknown length
 * 94854 142336: contig of 47483 bp in length
 * 142337 142437: gap of unknown length
 * 142437 173071: contig of 30635 bp in length.

FEATURES

source
 1. .173071
 /organism="Oryctolagus cuniculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9986"
 /clone="LB1-138P6"
 /clone_lib="LB1"
 /note="BAC resource: http://bacpac.chori.org/
 breed: New Zealand White"
 1. .17946
 /note="clone overlaps with GenBank Accession Number
 AC163953 clone LB1-24f21 (center project name izw)"
 misc_feature
 1. .2231
 /note="assembly_fragment
 clone_end:T7
 vector_side:left"
 gap
 2232. .2331
 /estimated_length=unknown
 misc_feature
 2332. .6543
 /note="assembly_fragment"
 gap
 6544. .6643
 /estimated_length=unknown
 misc_feature
 6644. .94753
 /note="assembly_fragment"
 gap
 94754. .94853
 /estimated_length=unknown
 misc_feature
 94854. .142336
 /note="assembly_fragment"
 gap
 142337. .142436
 /estimated_length=unknown
 misc_feature
 142437. .173071
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"

ORIGIN

Query Match 53.8%; Score 22.6; DB 12; Length 173071;
 Best Local Similarity 75.7%; Pred. No. 5.7e+02;
 Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 GCCCAGCTTGCCGAGCCCTTACAAGACTATGAAG 37
 |||||
 Db 77899 GCCCTGCCTTCCCTAGCCCAAGACCATGAAG 77863
 |||||
 RESULT 69
 AC091441/c 188622 bp DNA linear HTG 04-JUN-2001
 LOCUS
 DEFINITION Homo sapiens chromosome RPC1-11 clone RP11-693B19, WORKING DRAFT

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

SEQUENCE, 40 unordered pieces.
 AC091441
 AC091441.3 GI:14210637
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 Homo sapiens (human)
 Homo sapiens
 Mammalia; Eutelestomi; Chordata; Vertebrata; Eutelestomi;
 Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
 Homnidae; Homo.
 1 (bases 1 to 188622)
 Waterston,R.H.
 The sequence of Homo sapiens clone
 Unpublished
 2 (bases 1 to 188622)
 Waterston,R.H.
 Direct Submission
 Submitted (20-APR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On May 27, 2001 this sequence version replaced gi:13775308.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H_NH0693B19
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Chemistry: Dye-terminator Big Dye; 41% of reads
 Chemistry: Dye-terminator Big Dye; 41% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 160140 bases at least Q40
 Consensus quality: 168580 bases at least Q30
 Consensus quality: 173999 bases at least Q20
 Insert size: 179000; agarose-fp
 Insert size: 184722; sum-of-contigs
 Quality coverage: 3.11 in Q20 bases; agarose-fp
 Quality coverage: 3.10 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 40 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1660: contig of 1660 bp in length
 * 1661 1760: gap of unknown length
 * 1761 2764: contig of 1004 bp in length
 * 2765 2864: gap of unknown length
 * 2865 4001: contig of 1137 bp in length
 * 4002 4101: gap of unknown length
 * 4102 6019: contig of 1918 bp in length
 * 6020 6120: gap of unknown length
 * 6120 7712: contig of 1593 bp in length
 * 7713 7812: gap of unknown length
 * 7813 9594: contig of 1782 bp in length
 * 9595 9694: gap of unknown length
 * 9695 11105: contig of 1411 bp in length
 * 11106 11205: gap of unknown length
 * 11206 12957: contig of 1752 bp in length
 * 12958 13057: gap of unknown length
 * 13058 15321: contig of 2264 bp in length
 * 15322 15421: gap of unknown length
 * 15422 17637: contig of 2216 bp in length
 * 17638 17737: gap of unknown length
 * 17738 20175: contig of 2438 bp in length
 * 20176 20275: gap of unknown length
 * 20276 22783: contig of 2508 bp in length

* 22784 22883: gap of unknown length
* 22884 25416: contig of 2533 bp in length
* 25417 25516: gap of unknown length
* 25517 28197: contig of 2681 bp in length
* 28198 28297: gap of unknown length
* 28298 30593: contig of 2236 bp in length
* 30594 30693: gap of unknown length
* 30694 33550: contig of 2857 bp in length
* 33551 37481: contig of 3831 bp in length
* 37482 37581: gap of unknown length
* 37582 39009: contig of 2328 bp in length
* 39010 40009: gap of unknown length
* 40010 42913: contig of 2904 bp in length
* 42914 43013: gap of unknown length
* 43014 46368: contig of 3355 bp in length
* 46369 46468: gap of unknown length
* 46469 51176: contig of 4708 bp in length
* 51177 51276: gap of unknown length
* 51277 54607: contig of 3331 bp in length
* 54608 54707: gap of unknown length
* 54708 59400: contig of 4693 bp in length
* 59401 64255: contig of 4755 bp in length
* 64256 64355: gap of unknown length
* 64356 69910: contig of 5555 bp in length
* 69911 70010: gap of unknown length
* 70011 75918: contig of 5908 bp in length
* 75919 76018: gap of unknown length
* 76019 83100: contig of 7082 bp in length
* 83101 83200: gap of unknown length
* 83201 87479: contig of 4279 bp in length
* 87480 87579: gap of unknown length
* 87580 93930: contig of 6351 bp in length
* 93931 94030: gap of unknown length
* 94031 99775: contig of 5745 bp in length
* 99776 99875: gap of unknown length
* 99876 105199: contig of 5324 bp in length
* 105200 105299: gap of unknown length
* 105300 111500: contig of 6201 bp in length
* 111501 111600: gap of unknown length
* 111601 120914: contig of 9314 bp in length
* 120915 120914: gap of unknown length
* 121015 128821: contig of 7807 bp in length
* 128822 128921: gap of unknown length
* 128922 138027: contig of 9106 bp in length
* 138028 138127: gap of unknown length
* 138128 147606: contig of 9479 bp in length
* 147607 147606: gap of unknown length
* 147707 153904: contig of 6198 bp in length
* 153905 154004: gap of unknown length
* 154005 164023: contig of 10019 bp in length
* 164024 164123: gap of unknown length
* 164124 174952: contig of 10829 bp in length
* 174953 175052: gap of unknown length
* 175053 188622: contig of 13570 bp in length.

FEATURES

source
1..188622
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="RPC1-11"
/clone="RP11-693B19"
1..1660
/note="assembly_name:Contig10"
1661..1760
/estimated_length=unknown
1761..2764
/note="assembly_name:Contig11"
2765..2864
/estimated_length=unknown
2865..4001
/note="assembly_name:Contig17"

misc_feature
gap
misc_feature
gap
misc_feature

gap 4002..4101
/estimated_length=unknown
misc_feature 4102..6019
/note="assembly_name:Contig18"
gap 6020..6119
/estimated_length=unknown
misc_feature 6120..7712
/note="assembly_name:Contig19"
gap 7713..7812
/estimated_length=unknown
misc_feature 7813..9594
/note="assembly_name:Contig20"
gap 9595..9694
/estimated_length=unknown
misc_feature 9695..11105
/note="assembly_name:Contig21"
gap 11106..11205
/estimated_length=unknown
misc_feature 11206..12957
/note="assembly_name:Contig22"
gap 12958..13057
/estimated_length=unknown
misc_feature 13058..15321
/note="assembly_name:Contig23"
gap 15322..15421
/estimated_length=unknown
misc_feature 15422..17637
/note="assembly_name:Contig25"
gap 17638..17737
/estimated_length=unknown
misc_feature 17738..20175
/note="assembly_name:Contig26"
gap 20176..20275
/estimated_length=unknown
misc_feature 20276..22783
/note="assembly_name:Contig27"
gap 22784..22883
/estimated_length=unknown
misc_feature 22884..25416
/note="assembly_name:Contig28"
gap 25417..25516
/estimated_length=unknown
misc_feature 25517..28197
/note="assembly_name:Contig29"
gap 28198..28297
/estimated_length=unknown
misc_feature 28298..30593
/note="assembly_name:Contig30"
gap 30594..30693
/estimated_length=unknown
misc_feature 30694..33550
/note="assembly_name:Contig31"

Query Match 53.8%; Score 22.6; DB 12; Length 188622;
Best Local Similarity 75.7%; Pred. No. 5.8e+02;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTA 39
|||||
Db 75903 CCACACCTGGCCAGCCCTTACAAATTTTGTGAATA 75867

RESULT 70
DQ211932/c
LOCUS DQ211932 1857 bp mRNA linear INV 11-OCT-2005
DEFINITION Toxoplasma gondii mitochondrial alternative NADH dehydrogenase 1
(anhd-1) mRNA, complete cds; nuclear gene for mitochondrial
product.
ACCESSION DQ211932
VERSION DQ211932.1 GI:76561723
KEYWORDS
SOURCE Toxoplasma gondii
ORGANISM Toxoplasma gondii

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
1 (bases 1 to 1857)
AUTHORS
TITLE
The sequence of the respiratory chain components of Toxoplasma
gondii
Unpublished
JOURNAL
2 (bases 1 to 1857)
AUTHORS
TITLE
Direct Submission
Submitted (19-SEP-2005) Dept. of Bacteriology, Inst. of Medical
Microbiology, Kreuzberg 57, Goettingen, D 37075, Germany
Location/Qualifiers
1. .1857
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="RH"
/db_xref="taxon:5811"
/dev_stage="tachyzoite"
1. .1857
/gene="andh-1"
1. .1857
/gene="andh-1"
/EC_number="1.6.5.3"
/function="NADH:ubiquinone oxidoreductase"
/note="non-proton-pumping, rotenone-resistant single
polypeptide; belongs to the Pyr_redox (pyridine
nucleotide-disulphide oxidoreductase) family; localized to
the internal membrane of the mitochondrion"
/codon_start=1
/product="mitochondrial alternative NADH dehydrogenase 1"
/protein_id="ABA44355.1"
/db_xref="GI:76561724"
/translation="MAGQWLLLAGSVPMLSLPARCDSPASPSPPSKRVSALLA
QPFRGFAFSPVSTWMTKTSFRLRTGLSPAAVASAVASPAAREAPARRQKV
VVGSGADLVADLDMTRFVSPEDYFTPLPSPVCGTLPASACMTGVRELL
VGGVPCGSGFVEGRVAETCTPKKVRQSTHGKAQDAREWESYDYLVAAGADVNTF
GVGPKENAFVKELEDARRLSALFDVETASVPGVSEEEKKLHFVVVGAAGPTGV
EVAERIDDFQAGATGHPQLRPLVRITVVEMLPTVLAAYNNDVQAPKELLEENPRV
DLCLSQSVGVGPDSPVKTRAGQGVVEKEMPCGLLWASGKSPKVCULDARKTAE
LREAOQSPVILVDQMKVRGCEVYALGCCRLSPPLVQHADTLTYEATATGAAT
DWLEAREPKLSTVPFLQASSYDYFSQKPRQTMKEQFVKLLADIDAAYPAPTAQN
AKQAGRYLAQTAFNPSPVEEKRRAPAFVNTQTRGALVYLGHGQAAADIEGWRTPLGGA
TLLWKAAYLQWLTHNAVACLGWLRSLVGRVACREHLDTGETVYVDRRK"

Query Match 53.3%; Score 22.4; DB 13; Length 1857;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 3 CCACGCTGGCCGAGCCCTTACAAAGACTATGAGTAAG 42
|||||
DB 829 CCACGCGCTGCGAGCCCGCTCCACGACGACGAAGTGAAG 790
|||||

RESULT 71
AC024754/c
LOCUS
DEFINITION
AC024754
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 11959)
WormBase Consortium
Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
Science 282 (5396). 2012-2018 (1998)

9851916
PUBMED
REFERENCE
AUTHORS
Bemis, G.
TITLE
The sequence of C. elegans cosmid Y32G9B
Unpublished (2001)
JOURNAL
3 (bases 1 to 11959)
AUTHORS
TITLE
Waterston, R.H.
Direct Submission
Submitted (01-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 11959)
AUTHORS
TITLE
Waterston, R.
Direct Submission
Submitted (24-MAR-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 11959)
AUTHORS
TITLE
Waterston, R.
Direct Submission
Submitted (06-JUL-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
6 (bases 1 to 11959)
AUTHORS
TITLE
Waterston, R.
Direct Submission
Submitted (24-MAY-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
7 (bases 1 to 11959)
AUTHORS
TITLE
Waterston, R.
Direct Submission
Submitted (21-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
8 (bases 1 to 11959)
AUTHORS
TITLE
WormBase Consortium
Direct Submission
Submitted (22-SEP-2004) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
On Jul 6, 2001 this sequence version replaced gi:7140306.
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate chemistry
or covered by high quality data (i.e., phred quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by sequence from
more than one m13 subclone.
For a graphical representation of this clone sequence and its
analysis see:
http://www.wormbase.org/db/seq/sequence?name=Y32G9B;class=Sequence
NEIGHBORING CLONE INFORMATION
The 5' clone is Y46H3A, 400 bp overlap; the 3' clone is T10B5, 600
bp overlap.
NOTES:

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES

Location/Qualifiers

1. .34483
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP13-120P16"

misc_feature

1. .2005
/notes="overlaps bases 119624..121628 of clone AC117517"

repeat_region

127. .200
/function="clone overlap"

repeat_region

466. 486
/rpt_family="L1PA16"

repeat_region

945. 1025
/rpt_family="(T)n"

repeat_region

complement(1288..1571)
/rpt_family="L1PA16"

repeat_region

complement(1920..2014)
/rpt_family="AluY"

repeat_region

2118. 2157
/rpt_family="L2"

repeat_region

complement(2694..3316)
/rpt_family="(TC)n"

repeat_region

3749. 3895
/rpt_family="L1ME3A"

repeat_region

complement(4057..4240)
/rpt_family="MER58A"

repeat_region

4243. 4491
/standard_name="STS-T47077"

repeat_region

4264. 4371
/rpt_family="L2"

repeat_region

5068. 5372
/rpt_family="AluYb8"

repeat_region

5373. 5395
/rpt_family="(CAA)n"

repeat_region

5412. 5563
/rpt_family="MIR"

repeat_region

5591. 5614
/rpt_family="AT-rich"

repeat_region

5644. 5716
/rpt_family="CT-rich"

repeat_region

5921. 5950
/rpt_family="AT-rich"

repeat_region

6072. 6500
/rpt_family="MLT1J"

repeat_region

complement(6886..6898)
/rpt_family="MIR"

repeat_region

7236. 7282
/rpt_family="(TCTA)n"

repeat_region

complement(7284..7565)
/rpt_family="AluX"

repeat_region

complement(7672..8117)
/rpt_family="MER49"

repeat_region

complement(8174..8381)
/rpt_family="LTR59"

repeat_region

complement(8420..8719)
/rpt_family="AluY"

repeat_region

9116. 9421

repeat_region
/rpt_family="AluSg"
complement(9656..10044)
/rpt_family="L1"
10180. 10201
/rpt_family="AT-rich"
complement(10202..10316)
/rpt_family="L1MA9"
complement(10552..10646)
/rpt_family="MIR"
complement(10778..10908)
/rpt_family="MIR"
12280. 12301
/rpt_family="(T)n"
complement(12473..12801)
/rpt_family="L2"
complement(12824..13099)
/rpt_family="LTR16C"
complement(13125..13464)
/rpt_family="L2"
13471. 13634
/rpt_family="MER63A"
complement(13858..14026)
/rpt_family="MIR"
complement(14033..14097)
/rpt_family="L1MC5"
complement(14116..14302)
/rpt_family="L1MC/D"
14310. 15141
/rpt_family="L1PA7"
complement(15182..15328)
/rpt_family="L1MB8"
complement(15385..15432)
/rpt_family="L1MC4"
15521. 15696
/rpt_family="MER5B"
16163. 16317
/rpt_family="MER5A"
16294. 16425
STS

Query Match 53.3%; Score 22.4; DB 5; Length 34483;
Best Local Similarity 72.5%; Pred. No. 4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 41

Db 12445 CCACACCTGGCATGCCATGACAAATACGATCAAGTTAA 12406

RESULT 73

CR522870_21/c

WPCOMMENT

Sequence split into 36 fragments LOCUS CR522870 Accession CR522870

Fragment Name	Begin	End
CR522870_00	1	110000
CR522870_01	100001	210000
CR522870_02	200001	310000
CR522870_03	300001	410000
CR522870_04	400001	510000
CR522870_05	500001	610000
CR522870_06	600001	710000
CR522870_07	700001	810000
CR522870_08	800001	910000
CR522870_09	900001	1010000
CR522870_10	1000001	1110000
CR522870_11	1100001	1210000
CR522870_12	1200001	1310000
CR522870_13	1300001	1410000
CR522870_14	1400001	1510000
CR522870_15	1500001	1610000
CR522870_16	1600001	1710000
CR522870_17	1700001	1810000
CR522870_18	1800001	1910000
CR522870_19	1900001	2010000

RBS
CDS
YTAELRWGREGRRVIAIVOTDPTAPGKAQALNVALMVVSDWIVPESGYBAKTEALL
REQRRFLKPMREDAQSQVDFDWMIDIGAGRELPMVEYGRHDPKYLARKQAKCAYI
TOHYPTGWASHDASADPRGMNIPPPPIRHP"
4559. .4565
/note="ORF7"
4575. .5309
/note="ORF7"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAC86742.1"
/db_xref="GI:18077106"
/db_xref="GOA:Q8VMP7"
/db_xref="UniProtKB/TrEMBL:O8VMP7"
/db_xref="InterPro:IPR005480"
/translation="MTVNWELVDSKLGKQVLELRLIGSTPHRRRRRRLQGIAPWS
VQKTEPYKHLGYTTDRAIAQCQGVSKSVKAYRESLGLIPFVPRKQALPLNHD
LRPKPLFGVFDQIEAQAGLAVQDIREALGFERVPVPGSSPQLIADYHGPL
LGYESLSTMTPAKISRHVGVPSYIILORCKDLGTPQTRVSVASRYDHLGKVPDAL
VARLAGVSRASIGARRQRLGIASATR"
complement(5405. .6613)
/gene="ruvB"
complement(5405. .6613)
/gene="ruvB"
/codon_start=1
/transl_table=11
/protein_id="CAC86743.1"
/db_xref="GI:18077107"
/db_xref="GOA:Q8VMP6"
/db_xref="UniProtKB/TrEMBL:O8VMP6"
/translation="MCSKDRPIFGLVDVNSFYCSERI FRPELRHRPVVLSNSDGC
IVATNEKALGLKMGDPYFKVNFLEQNGUAVFSSNTLYGELSHRVGVIAISLVE
WEQSIDELFVNLGLPEFIIIDVGRAIQARVLQWALTALPVGVIGTIGTKLAKAAQHASK
VREKTEPYKHLGYTTDRAIAQCQGVSKSVKAYRESLGLIPFVPRKQALPLNHD
LRPKPLFGVFDQIEAQAGLAVQDIREALGFERVPVPGSSPQLIADYHGPL
LGYESLSTMTPAKISRHVGVPSYIILORCKDLGTPQTRVSVASRYDHLGKVPDAL
VARLAGVSRASIGARRQRLGIASATR"
complement(5405. .6613)
/gene="ruvB"
complement(5405. .6613)
/gene="ruvB"
/codon_start=1
/transl_table=11
/protein_id="CAC86744.1"
/db_xref="GI:18077108"
/db_xref="GOA:Q8VMP5"
/db_xref="InterPro:IPR006197"
/db_xref="UniProtKB/TrEMBL:O8VMP5"
/translation="MNVVSIQFLRDSRQVLAVLAGSAGGFPSPAADYERPPISLDEL
LNIRAAHVFIYRVGESMRDAGIFDGVLIIVDRSITPCSGQIVLAYVDNQPLVKRLAK
APGDCWLLSANPAYKTTTPDEYASVDVFGVVTWDLTKHF"
complement(6625. .6630)
/gene="ruvB"
complement(6625. .6630)
/gene="ruvB"
complement(7039. .7046)
/gene="ruvA"
7158. .7163
/note="ORF11"
7172. .8122
/note="ORF11"
/codon_start=1
/transl_table=11
/product="putative ligase"
/protein_id="CAC86745.1"
/db_xref="GI:18077109"
/db_xref="GOA:Q8VMP4"
/db_xref="InterPro:IPR000977"
/db_xref="UniProtKB/TrEMBL:O8VMP4"
/translation="MAVKKSQLPARIQPOLVALLTRPPAGWAYEIKFDGYRMFARID

RBS
CDS
AGVGLTKNGYDWTARMRLAODLRSLPIRGAWLDGEVIVODVDRPAFOALQAEFAT
GRTHLVFAFDLLFIENVDRSRVERRELRLVLEQVDLDQVYSYDLDADPSSL
LASACALGIEGIVGKRLGSVAGEDGSKWLKVCVORQEFIIIGFTRSSAGISLLIG
LHNDGQQLQYAGRVRSRGTFRRELDRLKERLRLVTRITPAMEAPPKLSGAVVAEPVLV
CEVXPALPTPAGKVRHAYVIALREDKPSAISLESSTDPV"
8142. .8147
/note="ORF12"
8153. .8428
/note="ORF12"
/codon_start=1

Query Match 53.3%; Score 22.4; DB 15; Length 116580;
Best Local Similarity 72.5%; Pred. No. 6.1e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCCGAGCCCTTACAAAGACTATGAAGTAA 40
Db 98830 GCGCCGCATGACCGCGCGTGTACCAAGAGCATGGAGAA 98869

RESULT 75
AC068600/c
LOCUS
DEFINITION Homo sapiens BAC clone RP11-567N20 from 2, complete sequence.
ACCESSION AC068600
VERSION AC068600.5 GI:15341622
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 136657)
AUTHORS Abbott.A., Spalding.L., Phillips.A., Elliott.G. and Grewal.N.
TITLE The sequence of Homo sapiens BAC clone RP11-567N20
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 136657)
AUTHORS Waterston.R.H.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 136657)
AUTHORS Waterston.R.H.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 136657)
AUTHORS Waterston.R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 136657)
AUTHORS Wilson.R.K.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Aug 29, 2001 this sequence version replaced gi:15144353.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics

Center project name: H_NH0567N20

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Mes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPEC11 human BAC library was made from the blood of one male donor, as described by Oseogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-249F13, 2000 bp overlap; the clone sequenced to the right is RP11-16D24. Actual start of this clone is at base position 149708 of RP11-249F13; actual end is at base position 136657 of RP11-567N20.

Data from AC011936 was used to finish the clone, AC068600. Polymorphisms have been identified between AC011936 and AC068600.

FEATURES

source
Location/Qualifiers
1..136657
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-567N20"
/clone_lib="RFC1-11"

ORIGIN

Query Match 53.3%; Score 22.4; DB 5; Length 136657;
Best Local Similarity 72.5%; Pred No. 6.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCGAGCCCTTACAAAGACTATCAAGTAAAG 42
Db 59046 CAAAACTAGACCAAGGCTTTACAAAACCTAAGAGTATAG 59007

RESULT 76
AC006732 159453 bp DNA linear HTG 23-FEB-1999
LOCUS Caenorhabditis elegans clone Y32G9, *** SEQUENCING IN PROGRESS ***,
DEFINITION 9 unordered pieces.
ACCESSION AC006732
VERSION AC006732.1 GI:4263206
KEYWORDS HTG; HTGS PHASE1.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 159453)
AUTHORS Waterston, R.H.
JOURNAL The sequence of Caenorhabditis elegans clone
TITLE Unpublished
REFERENCE 2 (bases 1 to 159453)
AUTHORS Waterston, R.H.
JOURNAL Direct Submission
TITLE Submitted (23-FEB-1999) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT * NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2042: contig of 2042 bp in length
2043 2056: gap of unknown length
2057 5001: contig of 2945 bp in length
5002 5015: gap of unknown length
5016 11580: contig of 6565 bp in length
11581 11594: gap of unknown length
11595 20569: contig of 8975 bp in length
20570 20583: gap of unknown length
20584 42974: contig of 22391 bp in length
42975 42988: gap of unknown length
42989 59017: contig of 16029 bp in length
59018 59031: gap of unknown length
59032 88258: contig of 29227 bp in length
88259 88272: gap of unknown length
88273 117204: contig of 28932 bp in length
117205 117218: gap of unknown length
117219 159453: contig of 42235 bp in length.

FEATURES

source
Location/Qualifiers
1..159453
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/db_xref="taxon:6239"
/clone="Y32G9"
2043..2056
/estimated_length=unknown
5002..5015
/estimated_length=unknown
11581..11594
/estimated_length=unknown
20570..20583
/estimated_length=unknown
42975..42988
/estimated_length=unknown
59018..59031
/estimated_length=unknown
88259..88272
/estimated_length=unknown
117205..117218
/estimated_length=unknown

ORIGIN

Query Match 53.3%; Score 22.4; DB 12; Length 159453;
Best Local Similarity 81.2%; Pred No. 6.8e+02;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTA 32
Db 21125 GCTCACCTTGGCCTTAGGCTTAACAAAGACTA 21156

RESULT 77
AC025518 176355 bp DNA linear PRI 29-JUN-2002
LOCUS Homo sapiens chromosome 17, clone RP11-713H12, complete sequence.
DEFINITION AC025518
ACCESSION AC025518
VERSION AC025518.8 GI:21629418
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 176355)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176355)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepe, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehouck, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivar, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL

Submitted (09-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 176355)

REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazar, B., Choepe, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehouck, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (15-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 176355)

REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazar, B., Choepe, Y., Collamore, A., Cooke, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

Submitted (29-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 29, 2002 this sequence version replaced gi:21427869.

COMMENT

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8285
Center clone name: 713_H_12

FEATURES

Source	Location/Qualifiers
1..176355	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosomes="17"
	/map="17"
	/clone="RP11-713H12"
	/clone_lib="RP11-11 Human Male BAC"
	/complement(30..88)
	/rpt_family="L3"
89..388	/rpt_family="AluY"
	/complement(389..503)
	/rpt_family="L3"
591..640	/rpt_family="L2"
1380..2441	/rpt_family="L2"
	/complement(2781..2876)
	/rpt_family="L2"
3360..3655	/rpt_family="AluJb"
4230..4380	/rpt_family="AluY"
4449..4745	/rpt_family="AluSg"
4931..4987	/rpt_family="TAAA"n"
5006..5074	/rpt_family="TTTC"n"
	/complement(5183..5252)
	/rpt_family="HAL1"
5680..5735	/rpt_family="CT-rich"
5915..6212	/rpt_family="AluSx"
6636..6943	/rpt_family="AluSp"
	/complement(7166..7384)
	/rpt_family="LIME1"
	/complement(7385..7732)
	/rpt_family="MER2"
	/complement(7733..8912)
	/rpt_family="LIME1"
8913..8940	/rpt_family="TTTG"n"
	/complement(8941..9394)
	/rpt_family="LIME1"
9395..9432	/rpt_family="TA"n"
	/complement(9433..10165)
	/rpt_family="LIME1"
	/complement(10889..10989)
	/rpt_family="MIR"
11195..11492	/rpt_family="AluY"
11755..12059	/rpt_family="AluJb"
12354..12497	/rpt_family="MIR"

```

repeat_region complement(12811..12979)
/rpt family="MERSA"
repeat_region complement(13572..13634)
/rpt family="L2"
repeat_region complement(13667..14961)
/rpt family="L2"
repeat_region complement(15020..15149)
/rpt family="MLTIH"
repeat_region complement(15837..15949)
/rpt family="L1MC4"
repeat_region complement(15990..16091)
/rpt family="MER34"
repeat_region complement(16528..16681)
/rpt family="L1ME3"
repeat_region 16953..17049
/rpt family="LTR49"
repeat_region 17060..17202
/rpt family="TA)n"
repeat_region 17224..17534
/rpt family="Alu9p"
repeat_region 17609..18017
/rpt family="LTR56"
repeat_region 18788..18850
/rpt family="MIR"
repeat_region 18911..18947
/rpt family="MER39B"
repeat_region 20262..20301
/rpt family="CAT)n"
repeat_region complement(20318..20535)
/rpt family="MIR"
repeat_region complement(20626..20718)
/rpt family="MIR"
repeat_region 21147..21645
/rpt family="L2"
repeat_region complement(21797..21897)
/rpt family="L2"
repeat_region complement(22229..22388)
/rpt family="Alu9/x"
repeat_region complement(22822..22940)

Query Match 53.3%; Score 22.4; DB 5; Length 176355;
Best Local Similarity 72.5%; Pred.No.7e+02; 11; Indels 0; Gaps 0;
Matches 29; Conservative 0; Mismatches 0;

Qy 3 CCAGCTTGGCGAGCCCTACAAGACTATGAAGTAAAG 42
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59526 CCAGCGCTGGCTGAGCTCTCAGCAATTCATGCACTAAG 59565

RESULT 78
AC024640/c
LOCUS
DEFINITION Homo sapiens chromosome 2 clone RP11-75986 map 2, WORKING DRAFT
SEQUENCE 181914 bp DNA linear HTG 24-AUG-2002
ACCESSION AC024640
VERSION AC024640.3 GI:8076901
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 181914)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 2, clone RP11-75986
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 181914)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tsefaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 181914)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collamore,A., Cooke,P., DeArellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
McClean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tsefaye,S.,
Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7239651.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6847
Center Clone name: 759_B_6
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165475 bases at least Q40
Consensus quality: 173706 bases at least Q30
Consensus quality: 177163 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 179514; sum-of-contigs
Quality coverage: 3.8 in Q20 bases; agarose-fp
Quality coverage: 3.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

* as soon as it is available and the accession number will
* be preserved.

```

* 1412: contig of 1412 bp in length
* 1512: gap of 100 bp
* 2838: contig of 1326 bp in length
* 2839: gap of 100 bp
* 4092: contig of 1154 bp in length
* 4093: gap of 100 bp
* 7004: contig of 2812 bp in length
* 7104: gap of 100 bp
* 10208: contig of 3104 bp in length
* 10308: gap of 100 bp
* 12886: contig of 2578 bp in length
* 12987: gap of 100 bp
* 15356: contig of 2370 bp in length
* 15357: gap of 100 bp
* 19277: contig of 3821 bp in length
* 19377: gap of 100 bp
* 21875: contig of 2498 bp in length
* 21876: gap of 100 bp
* 25234: contig of 3259 bp in length
* 25334: gap of 100 bp
* 28758: contig of 3424 bp in length
* 28858: gap of 100 bp
* 33066: contig of 4748 bp in length
* 33706: gap of 100 bp
* 40285: contig of 6579 bp in length
* 40385: gap of 100 bp
* 46553: contig of 6168 bp in length
* 46554: gap of 100 bp
* 53439: contig of 6786 bp in length
* 53440: gap of 100 bp
* 63348: contig of 9809 bp in length
* 63448: gap of 100 bp
* 69266: contig of 5818 bp in length
* 69366: gap of 100 bp
* 79513: contig of 10147 bp in length
* 79514: gap of 100 bp
* 79613: contig of 10870 bp in length
* 90483: gap of 100 bp
* 90584: contig of 10895 bp in length
* 101478: gap of 100 bp
* 114101: contig of 12523 bp in length
* 114201: gap of 100 bp
* 12653: contig of 12452 bp in length
* 12654: gap of 100 bp
* 139164: contig of 12411 bp in length
* 139165: gap of 100 bp
* 159467: contig of 20203 bp in length
* 159468: gap of 100 bp
* 181914: contig of 22347 bp in length.

```

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-759B6"
/clone_lib="RPC1-11 Human Male BAC"
1..1412
/note="assembly_fragment"
1413..1512
/estimated_length=100
1513..2838
/note="assembly_fragment"
2839..2938
/estimated_length=100
2939..4092
/note="assembly_fragment"
clone end:SP6
vector side:right
4093..4192

```

misc_feature

gap

misc_feature

gap

misc_feature

gap

misc_feature

gap

misc_feature

gap

misc_feature

gap

misc_feature

gap

misc_feature

gap

misc_feature

Query Match 53.3%; Score 22.4; DB 12; Length 181914;
Best Local Similarity 72.5%; Pred. No. 7.1e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 CCACGCTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 133604 CAAAAGTACCAAGGCTTTACAAAACTAAGAATATAG 133565

RESULT 79

AC116170/c

LOCUS

DEFINITION

SEQUENCE, 4 unordered pieces.

AC116170.2

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 186870)

Birren,B., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 17, clone RP11-855E10

Unpublished

2 (bases 1 to 186870)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

/estimated_length=100

4193..7004

/note="assembly_fragment"

7005..7104

/estimated_length=100

7105..10208

/note="assembly_fragment"

10209..10308

/estimated_length=100

10309..12886

/note="assembly_fragment"

12887..12986

/estimated_length=100

12987..15356

/note="assembly_fragment"

15357..15457

/estimated_length=100

15457..19277

/note="assembly_fragment"

19278..19377

/estimated_length=100

19378..21875

/note="assembly_fragment"

21876..21975

/estimated_length=100

21976..25234

/note="assembly_fragment"

25235..25334

/estimated_length=100

25335..28758

/note="assembly_fragment"

28759..28858

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 186870)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 24, 2002 this sequence version replaced gi:19703298.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIPR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L26190

Center clone name: 855_E10

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 185544 bases at least Q40

Consensus quality: 185963 bases at least Q30

Consensus quality: 186196 bases at least Q20

Insert size: 188000; agarose-fp

Insert size: 186570; sum-of-contigs

Quality coverage: 17.0 in Q20 bases; agarose-fp

Quality coverage: 17.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

* 1 16661: contig of 16661 bp in length
 * 16662 16761: gap of 100 bp
 * 16762 42342: contig of 25581 bp in length
 * 42343 42442: gap of 100 bp
 * 42443 114057: contig of 71615 bp in length
 * 114058 114157: gap of 100 bp
 * 114158 186870: contig of 72713 bp in length.

FEATURES
source

Location/Qualifiers
 1. 186870
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-855E10"
 /clone_lib="RFCI-11 Human Male BAC"
 1. 16661
 /note="assembly_fragment"
 clone_end.SP6
 vector_side:left
 16662..16761
 /estimated_length=100
 16762..42342
 /note="assembly_fragment"
 42343..42442
 /estimated_length=100
 42443..114057
 /note="assembly_fragment"
 114058..114157
 /estimated_length=100
 114158..186870
 /note="assembly_fragment"

ORIGIN

Query Match 53.3%; Score 22.4; DB 12; Length 186870;
 Best Local Similarity 72.5%; Pred. No. 7.2e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 CCAGCTTGGCGGACCTTACAAAGACTATGAGTAAG 42

Db 87951 CCAGGCTTGGCTGAGCTCTCAGCAATTCATGCACTAAG 87912

RESULT 80
 AX400400
 LOCUS AX400400 594 bp DNA linear PAT 06-JUN-2002
 DEFINITION Sequence 76 from Patent WO0210453.
 ACCESSION AX400400
 VERSION AX400400.1 GI:21336580

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

REFERENCE
 AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and Elashoff,M.R.
 TITLE Molecular toxicology modeling
 JOURNAL Patent: WO 0210453-A 76 07-FEB-2002;
 Gene Logic, Inc. (US)

FEATURES

source
 1. 584
 /organism="Rattus norvegicus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10116"
 /note="EMBL/GenBank Accession No. AA818700"

ORIGIN

Query Match 52.9%; Score 22.2; DB 2; Length 584;
 Best Local Similarity 88.9%; Pred. No. 1.2e+02;
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAGTAAA 41

Db 373 GAGCCCTTACAAAGACTAAGACCAA 399

RESULT 81

AC104541

```

LOCUS      AC104541                      69203 bp      DNA      linear      HTG 13-DEC-2001
DEFINITION Mus musculus clone RP23-142P5, LOW-PASS SEQUENCE SAMPLING.
ACCESSION  AC104541
VERSION    AC104541.1  GI:17646941
KEYWORDS   HTG; HTGS PHASE0.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 69203)
AUTHORS    Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE      Mus musculus, clone RP23-142P5
JOURNAL    Unpublished
AUTHORS     2 (bases 1 to 69203)
           Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
           Anderson, S., Barna, N., Bastien, V., Boguslavsky, B., Boukhgalter, B.,
           Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
           Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
           Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
           Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
           Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
           Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
           Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
           Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
           MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
           McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
           Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
           Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
           Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
           Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
           Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
           Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
           Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
           Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
           Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
           Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19844
Center clone name: 142_P_5
-----
* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 764: contig of 764 bp in length
* 765 864: gap of 100 bp
* 865 1620: contig of 756 bp in length
* 1621 1720: gap of 100 bp
* 1721 2493: contig of 773 bp in length
* 2494 2593: gap of 100 bp
* 2594 3317: contig of 724 bp in length
* 3318 3417: gap of 100 bp
* 3418 4160: contig of 743 bp in length
* 4161 4260: gap of 100 bp
*
* 4261 5031: contig of 771 bp in length
* 5032 5131: gap of 100 bp
* 5132 5900: contig of 769 bp in length
* 5901 6000: gap of 100 bp
* 6001 6734: contig of 734 bp in length
* 6735 7598: contig of 764 bp in length
* 7599 7698: gap of 100 bp
* 7699 8457: contig of 759 bp in length
* 8458 8557: gap of 100 bp
* 8558 9286: contig of 729 bp in length
* 9287 9386: gap of 100 bp
* 9387 10132: contig of 746 bp in length
* 10133 10232: gap of 100 bp
* 10233 10899: contig of 757 bp in length
* 10899 11090: gap of 100 bp
* 11090 11870: contig of 781 bp in length
* 11871 11970: gap of 100 bp
* 11971 12752: contig of 782 bp in length
* 12753 12852: gap of 100 bp
* 12853 13311: contig of 779 bp in length
* 13312 13731: gap of 100 bp
* 13732 14493: contig of 762 bp in length
* 14494 14593: gap of 100 bp
* 14594 15316: contig of 723 bp in length
* 15317 15416: gap of 100 bp
* 15417 16135: contig of 719 bp in length
* 16136 16235: gap of 100 bp
* 16236 17003: contig of 768 bp in length
* 17004 17103: gap of 100 bp
* 17104 17877: contig of 774 bp in length
* 17878 17977: gap of 100 bp
* 17978 18077: contig of 730 bp in length
* 18078 18807: gap of 100 bp
* 18808 19548: contig of 741 bp in length
* 19549 19648: gap of 100 bp
* 19649 20362: contig of 714 bp in length
* 20363 20462: gap of 100 bp
* 20463 21221: contig of 759 bp in length
* 21222 21321: gap of 100 bp
* 21322 22056: contig of 735 bp in length
* 22057 22156: gap of 100 bp
* 22157 22926: contig of 770 bp in length
* 22927 23026: gap of 100 bp
* 23027 23793: contig of 767 bp in length
* 23794 23893: gap of 100 bp
* 23894 24634: contig of 741 bp in length
* 24635 24734: gap of 100 bp
* 24735 25495: contig of 761 bp in length
* 25496 25595: gap of 100 bp
* 25596 26356: contig of 761 bp in length
* 26357 26456: gap of 100 bp
* 26457 27197: contig of 741 bp in length
* 27198 27297: gap of 100 bp
* 27298 28039: contig of 742 bp in length
* 28040 28139: gap of 100 bp
* 28140 28887: contig of 748 bp in length
* 28888 28987: gap of 100 bp
* 28988 29760: contig of 773 bp in length
* 29761 29860: gap of 100 bp
* 29861 30636: contig of 776 bp in length
* 30637 30736: gap of 100 bp
* 30737 31500: contig of 764 bp in length
* 31501 31600: gap of 100 bp
* 31601 32358: contig of 758 bp in length
* 32359 32458: gap of 100 bp
* 32459 33229: contig of 771 bp in length
* 33230 33329: gap of 100 bp
* 33330 34052: contig of 723 bp in length
* 34053 34152: gap of 100 bp
* 34153 34911: contig of 759 bp in length
* 34912 35011: gap of 100 bp
* 35012 35744: contig of 733 bp in length

```

* 35745 35844: gap of 100 bp
* 35845 3615: contig of 771 bp in length
* 36715: gap of 100 bp
* 37453: contig of 738 bp in length
* 37454 37553: gap of 100 bp
* 37554 38275: contig of 722 bp in length
* 38276 38375: gap of 100 bp
* 38376 39097: contig of 722 bp in length
* 39098 39197: gap of 100 bp
* 39198 39943: contig of 746 bp in length
* 39944 40043: gap of 100 bp
* 40044 40816: contig of 773 bp in length
* 40817 40916: gap of 100 bp
* 41681: contig of 765 bp in length
* 41781: gap of 100 bp
* 41782 42545: contig of 764 bp in length
* 42546 42645: gap of 100 bp
* 42646 43395: contig of 750 bp in length
* 43396 44248: contig of 753 bp in length
* 44249 44348: gap of 100 bp
* 44349 45112: contig of 764 bp in length
* 45113 45212: gap of 100 bp
* 45213 45983: contig of 771 bp in length
* 45984 46083: gap of 100 bp
* 46084 46831: contig of 748 bp in length
* 46832 46931: gap of 100 bp
* 46933 47714: contig of 783 bp in length
* 47715 47814: gap of 100 bp
* 47815 48580: contig of 766 bp in length
* 48581 48680: gap of 100 bp
* 48681 49462: contig of 782 bp in length
* 49463 49562: gap of 100 bp
* 49563 50338: contig of 776 bp in length
* 50339 51159: contig of 721 bp in length
* 51160 51259: gap of 100 bp
* 51260 52017: contig of 758 bp in length
* 52018 52117: gap of 100 bp
* 52118 52883: contig of 766 bp in length
* 52884 52983: gap of 100 bp
* 52984 53752: contig of 769 bp in length
* 53753 53852: gap of 100 bp
* 53853 54593: contig of 741 bp in length
* 54594 54693: gap of 100 bp
* 54694 55434: contig of 741 bp in length
* 55435 55534: gap of 100 bp
* 55535 56261: contig of 727 bp in length
* 56262 56361: gap of 100 bp
* 56362 57100: contig of 739 bp in length
* 57101 57200: gap of 100 bp
* 57201 57969: contig of 769 bp in length
* 57970 58069: gap of 100 bp
* 58070 58840: contig of 771 bp in length

Query Match 52.9%; Score 22.2; DB 12; Length 69203;
Best Local Similarity 77.1%; Pred. No. 6.3e-02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
Db 53656 GCTTGGCAAGCTCTTAGAAGATGACATAAGTAAA 53690

RESULT 82
AC011600_3
WPCOMMENT
Sequence split into 4 fragments LOCUS AC011600 Accession AC011600
Fragment Name Begin End
AC011600_0 1 110000
AC011600_1 100001 210000
AC011600_2 200001 310000
AC011600_3 300001 373878

Continuation (4 of 4) of AC011600 from base 300001 (AC011600 Homo sapiens chromosome 12 c
Query Match 52.9%; Score 22.2; DB 12; Length 73878;
Best Local Similarity 77.1%; Pred. No. 6.4e-02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
Db 4908 GCCTTGACAGTTATTACAAAGATATGAAGTAAA 4942

RESULT 83
AC011600_2
WPCOMMENT
Sequence split into 4 fragments LOCUS AC011600 Accession AC011600
Fragment Name Begin End
AC011600_0 1 110000
AC011600_1 100001 210000
AC011600_2 200001 310000
AC011600_3 300001 373878

Continuation (3 of 4) of AC011600 from base 200001 (AC011600 Homo sapiens chromosome 12 c
Query Match 52.9%; Score 22.2; DB 12; Length 110000;
Best Local Similarity 77.1%; Pred. No. 7.4e-02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
Db 104908 GCCTTGACAGTTATTACAAAGATATGAAGTAAA 104942

RESULT 84
AC079597/c
LOCUS
DEFINITION Homo sapiens 156766 bp DNA linear PRI 25-SEP-2002
Human BAC Library) complete sequence.
ACCESSION AC079597.13 GI:13129401
VERSION AC079597.13
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 156766)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Honsi,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisseg,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marcondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzger,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,


```

/function="clone overlap"
complement(1..44)
/rpt_family="THE1C"
complement(45..399)
/rpt_family="THE1C-internal"
971..1076
/rpt_family="MIR"
2309..2364
/rpt_family="CATATA)n"
2365..2444
/rpt_family="MER47"
2493..2559
/rpt_family="MER47"
2569..2796

Query Match      52.9%; Score 22.2; DB 5; Length 156766;
Best Local Similarity 77.1%; Pred. No. 8.4e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 41
Db 70179 GCCTTGACAAGTTATTACAAAGAATATGAAGTAA 70145

RESULT 85
AC128161/c
LOCUS
DEFINITION AC128161 205274 bp DNA linear HTG 22-SEP-2002
Rattus norvegicus clone CH230-58G18, *** SEQUENCING IN PROGRESS
***, 5 unordered pieces.
ACCESSION AC128161.2 GI:23196340
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
1 (bases 1 to 205274)
Muzny,D,Marie., Metzker,M,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alebrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gunnaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Gebregeorgis,E., Gear,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Joliviet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackelneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
AUTHORS
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRAV
Center clone name: CH230-58G18
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 185782 bases at least Q40
Consensus quality: 188579 bases at least Q30
Consensus quality: 190025 bases at least Q20
Estimated insert size: 210634; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 13135: contig of 13135 bp in length
* 13136 13235: gap of unknown length
* 13236 191447: contig of 178212 bp in length
* 191448 191547: gap of unknown length
* 191548 192710: contig of 1163 bp in length
* 192711 192810: gap of unknown length
* 192811 192750: contig of 4440 bp in length
* 192751 197350: gap of unknown length
* 197351 205274: contig of 7924 bp in length.
Location/Qualifiers
1..205274

```

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-58G18"
13136..13235
/estimated_length=unknown
complement(152564..178037)
/note="clone_boundary"
clone end: r7
site: EcoRI
end sequence: BH302924
191448..191547
/estimated_length=unknown
192711..192810
/estimated_length=unknown
197251..197350
/estimated_length=unknown

gap
misc_feature

ORIGIN
Query Match      52.9%; Score 22.2; DB 12; Length 205274;
Best Local Similarity 88.9%; Pred. No. 9.2e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
Db 119277 GAGCCCTTACAAAGACTAAGAACCAA 119251
|||||
|||||

RESULT 86
AC171018/c
LOCUS
DEFINITION
AC171018
Gallus gallus BAC clone CH261-11F14 from chromosome ul, complete
sequence.
AC171018
VERSION
AC171018.2 GI:85362939
KEYWORDS
HTG.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 216886)
Schmidt, H. and Kozlowski, A.
The sequence of Gallus gallus BAC clone CH261-11F14
Unpublished (2001)
2 (bases 1 to 216886)
Wilson, R. K.
Direct Submission
Submitted (22-OCT-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 216886)
Wilson, R. K.
Direct Submission
Submitted (19-JAN-2006) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
On Jan 19, 2006 this sequence version replaced gi:78042466.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: J_AA011F14

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by

```

```

restriction digest.
SOURCE INFORMATION:
This sequence is the entire insert of the clone.
FEATURES
Source
1..216886
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/chromosome="ul"
/clone="CH261-11F14"
/clone_lib="CHORI261"
ORIGIN
Query Match      52.9%; Score 22.2; DB 11; Length 216886;
Best Local Similarity 77.1%; Pred. No. 9.4e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 8 CTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 91252 CTGTGCTGAGCACTTCCAAAGGATATGAAATAAAG 91218
|||||
|||||

RESULT 87
AC096068/c
LOCUS
DEFINITION
AC096068
Rattus norvegicus clone CH230-20F15, *** SEQUENCING IN PROGRESS
***, 4 unordered pieces.
AC096068
VERSION
AC096068.6 GI:24817916
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 226291)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huiyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensunewa, L., Loulsegged, H., Lozada, R. J., Lu, K., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackeleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,

```

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Snajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Hoit, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 226291)
 Worley, K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 226291)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 9, 2002 this sequence version replaced gi:22772219.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRIT
 Center clone name: CH230-20F15
 ----- Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 183145 bases at least Q40
 Consensus quality: 187692 bases at least Q30
 Consensus quality: 18935 bases at least Q20
 Estimated insert size: 184258; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 220799: contig of 220799 bp in length
 * 220800 220899: gap of unknown length
 * 221900 221990: contig of 1091 bp in length
 * 221991 222090: gap of unknown length
 * 222091 223397: contig of 1307 bp in length

* 223398 223497: gap of unknown length
 * 223498 226291: contig of 2794 bp in length.
 FEATURES
 source

1. 226291
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-20F15"

misc_feature

1. 835
 /note="clone_boundary
 clone end: T7
 site: EcoRI
 end sequence: BH319833"

misc_feature

174564..175853
 /note="wgs contig"

misc_feature

217818..220799
 /note="wgs contig"

gap

220800..220899
 /estimated_length=unknown

gap

221991..222090
 /estimated_length=unknown

gap

223398..223497
 /estimated_length=unknown

ORIGIN

Query Match 52.9%; Score 22.2; DB 12; Length 226291;
 Best Local Similarity 88.9%; Pred. No. 9.5e+02;
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
 |||||

Db 52562 GAGCCCTTACAAAGACTATGAAGTAAA 52536
 |||||

RESULT 88

AC126713 AC126713 247515 bp DNA linear HTG 09-MAY-2003
 LOCUS Rattus norvegicus clone CH230-5L14, WORKING DRAFT SEQUENCE, 6
 DEFINITION
 unorderd pieces.

ACCESSION

AC126713.4 GI:30466573

VERSION

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 247515)

AUTHORS

Murphy, D. Marie., Metzker, M. Lee., Abranzon, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, S., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshewa, L., Louised, H., Lozado, R.J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinkle, M., Strong, R., Sutton, A., Svatek, A., Tabor, J., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 247515)
 Worley, K.C.
 Direct Submission
 Submitted (09-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 AUTHORS
 JOURNAL

3 (bases 1 to 247515)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 9, 2003 this sequence version replaced gi:24941703.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GBK0
 Center clone name: CH230-5L14
 ----- Summary Statistics

Assembly program: Atlas;
 Consensus quality: 228033 bases at least Q40
 Consensus quality: 230259 bases at least Q30
 Consensus quality: 231440 bases at least Q20
 Estimated insert size: 237542; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 236249: contig of 236249 bp in length
 * 236250 236349: gap of unknown length
 * 236350 237425: contig of 1076 bp in length
 * 237426 237525: gap of unknown length
 * 237526 239390: contig of 1865 bp in length
 * 239391 239491: gap of unknown length
 * 239491 240713: contig of 1223 bp in length
 * 240714 240813: gap of unknown length
 * 240814 242869: contig of 2056 bp in length
 * 242870 242969: gap of unknown length
 * 242970 247515: contig of 4546 bp in length.

FEATURES

source
 1..247515
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-5L14"

misc_feature

1..1818
 /note="wgs_end_extension
 clone_end:T7"

misc_feature

12486..14214
 /note="wgs_end_extension
 clone_end:T7"

misc_feature

15239..15971
 /note="clone_boundary
 clone_end:T7"

misc_feature

end_sequence:BH362758"
 complement(233573..234468)
 /note="clone_boundary
 clone_end:Sp6"

gap

end_sequence:BH362760"
 236250..236349
 /estimated_length=unknown

gap

237426..237525
 /estimated_length=unknown
 239391..239490
 /estimated_length=unknown

gap

240714..240813
 /estimated_length=unknown
 242870..242969
 /estimated_length=unknown

gap

247515..247515
 /estimated_length=unknown

ORIGIN

Query Match 52.9%; Score 22.2; DB 12; Length 247515;
 Best Local Similarity 77.1%; Pred. No. 9.8e+02;
 Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAA 36
 Db 95194 CACAGGAGTGGCAGAGACCTACAGAGACTATGAA 95228

RESULT 89

AC122818

LOCUS

AC122818 Mus musculus BAC clone RP23-243016 from 17, complete sequence.

DEFINITION

AC122818 Mus musculus (house mouse)

ACCESSION

AC122818.4 GI:26801343

VERSION

HTG.

KEYWORDS

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 167735)

Trani, L., Kozlowicz, A. and Schatzkammer, K.

The sequence of Mus musculus BAC clone RP23-243016


```

repeat_region 11811..11899
/rpt family="B2"
repeat_region 11907..12036
/rpt family="Alu"
repeat_region 12118..12302
/rpt family="B2"
repeat_region 12319..12428
/rpt family="Alu"
repeat_region 12817..13011
/rpt family="B2"
repeat_region 13559..13891
/rpt family="ERVK"
repeat_region 13996..14381
/rpt family="RMER198"
repeat_region 15092..16198
/rpt family="ERVK"
repeat_region 16595..16887
/rpt family="ERVK"
repeat_region 17012..17191
/rpt family="B2"
repeat_region 17228..17404
/rpt family="L1"
repeat_region 17407..17466
/rpt family="B4"
repeat_region 18377..18803
/rpt family="ERVK"
repeat_region 18802..18885
/rpt family="ERVK"
repeat_region 18884..19056
/rpt family="ERVK"
repeat_region 19127..19448
/rpt family="ERVK"
repeat_region 19811..19900
/rpt family="B4"
repeat_region 19970..20101
/rpt family="Alu"
repeat_region 20294..20437
/rpt family="Alu"
repeat_region 20532..20721
/rpt family="B2"
repeat_region 20960..21390

Query Match 52.4%; Score 22; DB 6; Length 167735;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CCGAGCCCTTACAAGACTATGAGTAAAG 42
Db 160901 CCAAGCCAGTACAAGACAATGAGTAAAG 160930

RESULT 90
AC121302/c
LOCUS AC121302 190256 bp DNA linear ROD 28-JAN-2004
DEFINITION Mus musculus chromosome 13, clone RP23-163B12, complete sequence.
ACCESSION AC121302
VERSION AC121302.7 GI:41351614
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 190256)
Birren,B., Nussbaum,C. and Lander,E.
TITLE Mus musculus chromosome 13, clone RP23-163B12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 190256)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,

```

```

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lechoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
O'Neil,D., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 190256)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-NOV-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 190256)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-JAN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 28, 2004 this sequence version replaced gi:38564369.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

```

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24214
 Center clone name: 163_B12

FEATURES

```

source
1. .190256
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /chromosomes="13"
  /map="13"
  /clone="RP23-163B12"
  /clone_lib="RPCI-23 Female Mouse BAC"
1. .573
  /note="wgs end extension"
  /clone_end:SP6"
  complement(103. .872)
  /rpt_family="HAL1"
574. .579
  /note="clone boundary"
  /clone_end:SP6"
  site:ECORI"
  /rpt_family="B3A"
1994. .2163
  /rpt_family="B3A"
2218. .2500
  /rpt_family="RLTR9C"
2568. .2653
  /rpt_family="(TA)n"
2877. .2921
  /rpt_family="AT_rich"
4036. .4088
  /rpt_family="(TG)n"
9881. .10185
  /rpt_family="LIMA4A"
  complement(10613. .10714)
  /rpt_family="ETNERV"
  complement(13182. .13371)
  /rpt_family="B2_Mm1"
  complement(13888. .14073)
  /rpt_family="B4A"
14306. .14327
  /rpt_family="AT_rich"
  complement(14764. .14888)
  /rpt_family="MTE"
  complement(14922. .15099)
  /rpt_family="MTE"
15594. .15790
  /rpt_family="B3"
  complement(15825. .15989)
  /rpt_family="Lx9"
15987. .16183
  /rpt_family="Lx9"
16195. .16275
  /rpt_family="(TA)n"
16330. .16358
  /rpt_family="AT_rich"
  complement(18164. .18249)
  /rpt_family="RMER15"
  complement(18448. .18946)
  /rpt_family="Lx8"
20301. .20478
  /rpt_family="(TG)n"
20501. .20545
  /rpt_family="(TCTG)n"
20748. .20778
  /rpt_family="(A)n"
  complement(20844. .21038)
  /rpt_family="B4A"
21435. .21499

```

```

repeat_region
  /rpt_family="MIR"
21924. .21965
  /rpt_family="(T)n"
22031. .22052
  /rpt_family="AT_rich"
22799. .23516
  /rpt_family="L1_MM"
24908. .24937
  /rpt_family="AT_rich"
25101. .25224
  /rpt_family="RSINE1"
25452. .25499
  /rpt_family="AT_rich"
25600. .25658
  /rpt_family="(TA)n"
25840. .25880
  /rpt_family="AT_rich"
25980. .26081
  /rpt_family="A-rich"
26305. .26337
  /rpt_family="(TTCA)n"
26344. .26384
  /rpt_family="(TAAAA)n"
26424. .26468
  /rpt_family="B1_MM"
  complement(26469. .26703)
  /rpt_family="Lx6"
27377. .27424
  /rpt_family="(TAGA)n"
  complement(28896. .29877)
  /rpt_family="Lx2"
31058. .31242
  /rpt_family="(GAAA)n"
31334. .31462
  /rpt_family="B1_MM"

Query Match      52.4%; Score 22; DB 6; Length 190256;
Best Local Similarity 73.7%; Pred.No. 1.1e+03;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy      3  CCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 40
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      80046 CCACGCTTGGACAGACCCCTGAAAAAATATGATGAAA 80009

RESULT 91
AC122489/191710 bp DNA linear ROD 13-NOV-2003
LOCUS Mus musculus BAC clone RP24-386H10 from chromosome 13, complete
DEFINITION sequence.
ACCESSION AC122489
VERSION AC122489.3 GI:28316670
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 191710)
  Dignan,G. and Meyer,R.
  The sequence of Mus musculus BAC clone RP24-386H10
  Unpublished (2001)
  2 (bases 1 to 191710)
  Sequencing of Mus musculus
  Unpublished (2001)
  3 (bases 1 to 191710)
  McPherson,J.D. and Waterston,R.H.
  Direct Submission
  Submitted (01-JUN-2002) Genome Sequencing Center, 4444 Forest Park
  Parkway, St. Louis, MO 63108, USA
  4 (bases 1 to 191710)
  McPherson,J.D. and Waterston,R.H.
  Direct Submission

```

JOURNAL
 Submitted (12-FEB-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 191710)
 McPherson,J.D. and Waterston,R.H.
 Direct Submission
 Submitted (28-FEB-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 6 (bases 1 to 191710)
 Wilson,R.
 Direct Submission
 Submitted (13-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Feb 12, 2003 this sequence version replaced gi:22091363.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.edu
 ----- Summary Statistics
 Center project name: M_BB0386H10

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
 The RPCI-24 BAC Library has been constructed by Pieter de Jong and
 coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
 This sequence is the entire insert of the clone.

FEATURES

source	Location/Qualifiers
	1..191710
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="13"
	/map="13"
	/clone="RP24-386H10"
	/clone_lib="RPCI-24"
repeat_region	14..308
repeat_region	/rpt_family="MaLR"
repeat_region	312..393
repeat_region	/rpt_family="RMR198"
repeat_region	722..837
repeat_region	/rpt_family="Alu"
repeat_region	839..1094
repeat_region	/rpt_family="L1"
repeat_region	1355..1438
repeat_region	/rpt_family="L1"
repeat_region	1531..1712
repeat_region	/rpt_family="L1"
repeat_region	2281..2409
repeat_region	/rpt_family="B4"
repeat_region	3565..3639
repeat_region	/rpt_family="ID"
repeat_region	5003..5217
repeat_region	/rpt_family="B2"
repeat_region	6484..6554
repeat_region	/rpt_family="ID"
repeat_region	7011..7196
repeat_region	/rpt_family="L1"
repeat_region	7217..10174
repeat_region	/rpt_family="L1"
repeat_region	10172..10275
repeat_region	/rpt_family="L1"
repeat_region	10285..10402
repeat_region	/rpt_family="L1"
repeat_region	10406..10661
repeat_region	/rpt_family="L1"
repeat_region	10661..10743
repeat_region	/rpt_family="L1"
repeat_region	11004..12487
repeat_region	/rpt_family="L1"
repeat_region	12492..12769
repeat_region	/rpt_family="L1"
repeat_region	12770..13270
repeat_region	/rpt_family="ERVK"
repeat_region	13276..13525
repeat_region	/rpt_family="ERVK"
repeat_region	13635..13936
repeat_region	/rpt_family="ERVK"
repeat_region	15387..16293
repeat_region	/rpt_family="ERVK"
repeat_region	17018..17297
repeat_region	/rpt_family="ERVK"
repeat_region	17335..17438
repeat_region	/rpt_family="ERVK"
repeat_region	18297..18485
repeat_region	/rpt_family="ERVK"
repeat_region	18488..18989
repeat_region	/rpt_family="ERVK"
repeat_region	18990..19496
repeat_region	/rpt_family="L1"
repeat_region	20092..20336
repeat_region	/rpt_family="L1"
repeat_region	20635..20737
repeat_region	/rpt_family="B4"
repeat_region	22375..22975
repeat_region	/rpt_family="L1"
repeat_region	24214..24353
repeat_region	/rpt_family="Alu"
repeat_region	25054..25209
repeat_region	/rpt_family="ERVK"
repeat_region	26586..27654
repeat_region	/rpt_family="L1"
repeat_region	30689..31458
repeat_region	/rpt_family="L1"
repeat_region	32580..32749
repeat_region	/rpt_family="B2"
repeat_region	32801..33086
repeat_region	/rpt_family="ERVK"
repeat_region	37498..37706
repeat_region	/rpt_family="MaLR"
repeat_region	40488..40687
repeat_region	/rpt_family="L1"
repeat_region	41185..41319
repeat_region	/rpt_family="ERVK"
repeat_region	43768..43957
repeat_region	/rpt_family="B2"
repeat_region	44474..44659
repeat_region	/rpt_family="B4"
repeat_region	45350..45685
repeat_region	/rpt_family="MaLR"
repeat_region	46180..46376
repeat_region	/rpt_family="B2"
repeat_region	46411..46575
repeat_region	/rpt_family="L1"

```

repeat_region      46573..46825
                    /rpt_family="L1"
repeat_region      49034..49594
                    /rpt_family="L1"
repeat_region      50897..50945
                    /rpt_family="ERV1"
repeat_region      51121..51243
                    /rpt_family="B4"
repeat_region      51430..51624
                    /rpt_family="B4"
repeat_region      53385..54102
                    /rpt_family="L1"
repeat_region      55687..55810
                    /rpt_family="B4"
repeat_region      56725..56852
                    /rpt_family="L1"
repeat_region      56935..57004
                    /rpt_family="L1"
repeat_region      57010..57054
                    /rpt_family="Alu"
repeat_region      57058..57291
                    /rpt_family="L1"
repeat_region      59494..60476
                    /rpt_family="L1"
repeat_region      60475..60534
                    /rpt_family="L1"
repeat_region      61552..61785

```

Query Match 52.4%; Score 22; DB 6; Length 191710;

Best Local Similarity 73.7%; Pred. No. 1.1e+03;

Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Qy 3 CCACGCTTGGCGACCCCTTACAAAGACTATGAAGTAA 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110632 CCACGCTTGGACAGACCCCTGAAAAACTATGATGAAA 110595

```

RESULT 92

```

AC084383/c
LOCUS      AC084383      203589 bp      DNA      linear      ROD 31-OCT-2003
DEFINITION Mus musculus BAC clone RP23-10B20 from 17, complete sequence.
ACCESSION AC084383
VERSION   AC084383.1  GI:15778815
KEYWORDS  HTG.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 203589)
AUTHORS   Nguyen,C., Abbott,A., Elliott,G. and Hawkins,M.
TITLE     The sequence of Mus musculus BAC clone RP23-10B20
JOURNAL   Unpublished (2001)
REFERENCE 2 (bases 1 to 203589)
AUTHORS   Wilson,R.
TITLE     Sequencing of Mus musculus
JOURNAL   Unpublished (2001)
REFERENCE 3 (bases 1 to 203589)
AUTHORS   McPherson,J.D.
TITLE     Direct Submission
JOURNAL   Submitted (26-SEP-2001) Genetics, Genome Sequencing Center, 4444
           Forest Park Parkway, St. Louis, Missouri 63108, USA
REFERENCE 4 (bases 1 to 203589)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
           Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 203589)
AUTHORS   Wilson,R.
TITLE     Direct Submission
JOURNAL   Submitted (31-OCT-2003) Department of Genetics, Washington
           University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
-----
COMMENT

```

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 Center project name: M_BA0010B20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazuoto Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source	Location/Qualifiers
1..203589	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="17"
	/map="17"
	/clone="RP23-10B20"
	/clone_lib="RPCI-23"
46..184	/rpt_family="B4"
292..1382	/rpt_family="L1"
1432..1551	/rpt_family="Alu"
1591..1688	/rpt_family="B2"
1862..2137	/rpt_family="B4"
5906..5972	/rpt_family="Alu"
6678..6821	/rpt_family="Alu"
6940..7110	/note="Unresolved simple sequence repeat."
8021..8354	/rpt_family="MaLR"
8355..10211	/rpt_family="MaLR"
10216..10522	/rpt_family="MaLR"
10726..10844	/rpt_family="B4"
10837..10887	/rpt_family="Alu"
11250..11393	

```

/rpt_family="B4"
11606..11705
/rpt_family="B4"
12301..12496
/rpt_family="7SK"
13746..13826
/rpt_family="L2"
15416..15560
/rpt_family="Alu"
15634..15670
/rpt_family="L2"
15798..16052
/rpt_family="MaLR"
16183..16257
/rpt_family="L2"
16289..16570
/rpt_family="B4"
16594..16842
/rpt_family="L2"
18179..18363
/rpt_family="B2"
18371..18459
/rpt_family="L1"
18460..18599
/rpt_family="Alu"
18600..18744
/rpt_family="L1"
18745..18877
/rpt_family="Alu"
18984..19058
/rpt_family="L1"
19079..19131
/rpt_family="MaLR"
19132..19263
/rpt_family="MaLR"
19501..19851
/rpt_family="MaLR"
21000..21039
/rpt_family="L1"
21286..21439
/rpt_family="L1"
21783..22945
/rpt_family="L1"
23027..23224
/rpt_family="B2"
23505..23673
/rpt_family="L1"
23802..23913
/rpt_family="Alu"
24851..24904
/rpt_family="B4"
25986..26296
/rpt_family="MaLR"
27018..27185
/rpt_family="B2"
29142..29242
/rpt_family="Alu"
29508..29606
/rpt_family="Alu"
30880..31954
/notes="CpG Island (%GC=72.3, o/e=0.75, #CpGs=116)"
33292..33382
/rpt_family="Alu"
33736..33868
/rpt_family="Alu"
33885..33991
/rpt_family="B2"
34513..34648
/rpt_family="B4"
34799..34953
/rpt_family="B4"
37019..37152
/rpt_family="Alu"

```

```

repeat_region 37177..37261
/rpt_family="Alu"
41917..41995
/rpt_family="MIR"
43353..43602
/rpt_family="B4"
44128..44173
/rpt_family="B4"
44174..44526
/rpt_family="MaLR"
44527..44609
/rpt_family="B4"
48077..48211
/rpt_family="MaLR"
48220..48515
/rpt_family="MaLR"
49197..49315
/rpt_family="B4"
49244..49356
/rpt_family="Alu"

```

Query Match 52.4%; Score 22; DB 6; Length 203589;
 Best Local Similarity 83.3%; Pred. NO. 1.1e+03;
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 13 CCGAGCCCTTACAAAGACTATGAGGTAAAG 42
Db 200649 CCAAGCCAGTACAAAGACAATGAGGTAAAG 200620

```

RESULT 93

AC092090

LOCUS

DEFINITION

AC092090

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC092090 209153 bp DNA linear HTG 23-MAY-2002
 Canis familiaris clone RP81-300L5, WORKING DRAFT SEQUENCE, 8
 ordered pieces.
 AC092090
 HTG; HTGS PHASE2; HTGS_DRAFT.
 Canis familiaris (dog)
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.
 1 (bases 1 to 209153)
 Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
 Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
 Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Lalic,P.,
 Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
 Margulies,E.H., Masiello,C., Maskeri,B., Mastrian,S.D.,
 McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
 Portnoy,M.E., Prasad,A., Schueler,M.G., Stantropop,S., Thomas,J.W.,
 Thomas,P.J., Touchman,J.W., Turgeon,C., Vogt,J.L., Walker,M.A.,
 Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 209153)
 Green,E.D.
 Direct Submission
 Submitted (20-JUN-2001) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 209153)
 Green,E.D.
 Direct Submission
 Submitted (23-MAY-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On May 23, 2002 this sequence version replaced gi:14495384.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoonhgri.nih.gov
 ----- Project Information

Center project name: ayt
Center clone name: 300L05

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 207698 bases at least Q40
Consensus quality: 208254 bases at least Q30
Consensus quality: 208411 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 208453; sum-of-contigs
Quality coverage: 13.81x in Q20 bases; agarose-fp
Quality coverage: 11.13x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 28791: contig of 28791 bp in length
* 28792 28991: gap of unknown length
* 28892 94557: contig of 65666 bp in length
* 94558 121614: contig of 26957 bp in length
* 121615 121715: gap of unknown length
* 121715 177060: contig of 55346 bp in length
* 177061 177160: gap of unknown length
* 177161 180861: contig of 3701 bp in length
* 180862 180961: gap of unknown length
* 180962 189115: contig of 8254 bp in length
* 189116 200836: gap of unknown length
* 200837 200937: contig of 11521 bp in length
* 200937 209153: contig of 8217 bp in length.

FEATURES

source	AC111880	249159 bp	DNA	linear	HTG 15-NOV-2002
	Rattus norvegicus clone CH230-25K13, WORKING DRAFT SEQUENCE, 3 unordered pieces.				
ACCESSION	AC111880.5	GI:25006926			
VERSION	HTGS_PHASE1	HTGS DRAFT; HTGS_FULLTOP.			
KEYWORDS	Rattus norvegicus (Norway rat)				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 249159)				
AUTHORS	Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerr, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulsegged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,				
misc_feature	177061..177160	/note="assembly_fragment"			
	177161..180861	/estimated_length=unknown			
misc_feature	180862..180961	/note="assembly_fragment"			
	180962..189115	/estimated_length=unknown			
misc_feature	189116..209153	/note="assembly_fragment"			
	AC092091 clone RP81-312A9	GenBank Accession Number (center project name ayt)"			
gap	189116..189315				
	189316..200836	/estimated_length=unknown			
misc_feature	200837..200936	/note="assembly_fragment"			
	200937..209153	/estimated_length=unknown			
misc_feature	209154..209159	/note="assembly_fragment"			
	209160..209159	clone_end:SP6			
	209160..209159	vector_side:right"			
ORIGIN					
Query Match	52.4%	Score 22;	DB 12;	Length 209153;	
Best Local Similarity	73.7%	Pred. No. 1.1e+03;			
Matches 28;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;	
QY	4	CACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA	41		
Db	133200	CAGGCTTGGCTAGCTCTTCCAAAGACTCAGCTGAAA	133237		
RESULT 94					
AC111880/c					
LOCUS	AC111880	249159 bp	DNA	linear	HTG 15-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-25K13, WORKING DRAFT SEQUENCE, 3 unordered pieces.				
ACCESSION	AC111880.5	GI:25006926			
VERSION	HTGS_PHASE1	HTGS DRAFT; HTGS_FULLTOP.			
KEYWORDS	Rattus norvegicus (Norway rat)				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 249159)				
AUTHORS	Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerr, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulsegged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,				
misc_feature	1..28791	/note="assembly_fragment"			
	28792..28891	clone_end:R7			
	28892..28891	vector_side:left"			
gap	28892..28891	/estimated_length=unknown			
misc_feature	28892..94557	/note="assembly_fragment"			
	94558..94557	/estimated_length=unknown			
gap	94558..94557				
misc_feature	94558..121614	/note="assembly_fragment"			
	121615..121714	/estimated_length=unknown			
gap	121615..121714				
misc_feature	121715..177060	/estimated_length=unknown			

Manung,B., Mapua,P., Martin,K., Martin,R., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nwakolameh,O., Okwuonu,G., Olarnpunsgoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
 Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steidle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villaana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 249159)
 Worley,K.C.

Direct Submission
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 249159)

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 15, 2002 this sequence version replaced gi:23321716.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GOKK
 Center clone name: CH230-25K13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 239075 bases at least Q40
 Consensus quality: 242050 bases at least Q30
 Consensus quality: 243432 bases at least Q20
 Estimated insert size: 251750; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 245189: contig of 245189 bp in length
 * 245190: gap of unknown length
 * 245290 246792: contig of 1503 bp in length
 * 246793 246892: gap of unknown length
 * 246893 249159: contig of 2267 bp in length.

FEATURES
 source
 1..249159
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-25K13"
 1..1631
 /note="wgs_contig"
 242917..245189
 /note="wgs_contig"
 245190..245289
 /estimated_length=unknown
 246793..246892
 /estimated_length=unknown

ORIGIN
 Query Match 52.4%; Score 22; DB 12; Length 249159;
 Best Local Similarity 83.3%; Pred. No. 1.2e+03;
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 12 GCGAGCCCTTACAAAGACTATGAGTAAA 41
 |||||
 DB 210029 GCAGAGCCCTTATATGACTAAGACAAA 210000

AC127823 255568 bp DNA linear HTG 19-NOV-2002
 Rattus norvegicus clone CH230-265A10, WORKING DRAFT SEQUENCE, 4
 unordered pieces.

AC127823
 AC127823.3 GI:25077948
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
 1 (bases 1 to 255568)
 Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,C.,
 Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollings,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,


```

ORIGIN
      FKYRIEKKITNFVHKTLAHVPRIIERDGSIGAWASEGNES"

Query Match          51.9%;   Score 21.8;   DB 11;   Length 428;
Best Local Similarity 70.7%;   Pred. No. 1.6e+02;
Matches 29;   Conservative 0;   Mismatches 12;   Indels 0;   Gaps 0;

QY      1  GCCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAA 41
          |||||
DB      29  GGCCACCCCTTGACAAACACCTTAGAAGAAGATGAACCTTAA 69

RESULT 98
AY844389
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
Hyla arborea (European tree frog)

AY844389      428 bp      DNA      linear
Hyla arborea Rag-1 gene, partial cds.
AY844389
AY844389.1   GI:61725394

```

SOURCE: Hyla arborea (Batrachoseps) tree 11957

ORGANISM: Hyla arborea

AUTHORS: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylinae; Hyla.

REFERENCE: 1 (bases 1 to 428)

AUTHORS: Faiyovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A. and Wheeler, W.C.

TITLE: Systematic Review of the frog family Hylidae, with special reference to the Hylinae: Phylogenetic analysis and taxonomic revision

JOURNAL: Bull. Am. Mus. Nat. Hist. 294, 1-240 (2005)

REFERENCE: 2 (bases 1 to 428)

AUTHORS: Faiyovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A. and Wheeler, W.C.

TITLE: Direct Submission

JOURNAL: Submitted (01-DEC-2004) Herpetology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA

FEATURES: Location/Qualifiers

SOURCE: 1. .428

```

/organism="Hyla arborea"
/mol_type="genomic DNA"
/db_xref="taxon:121869"
/country="Germany: Petrade"
<1..>428
mRNA
/product="Rag-1"
<1..>428
CDS
/codon_start=3
/product="Rag-1"
/protein_id="AA54099.1"
/db_xref="GI:61725395"
/translation="PVEERKKQATLDKHLRKKMNLKPIRMNGNFARKLMKSTVEA
VCDLTPSEERQAALRELDLYLKMPVWRSSCPAKECPPELLQCYSVHSQRFALLSTYK
FKYRYEGKITNYFHKTLAHVPIIERDGSIGAWASEGNS"
ORIGIN
Query Match 51.9%; Score 21.8; DB 11; Length 428;
Best Local Similarity 70.7%; Pred. No. 1.6e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

QY	1	GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAA	41
Db	29	GGCAGCGCTTGACAAACACCTTTAGAAAGAAGATGAACCTTAA	69
RESULT 99			
AY844391	LOCUS		
AY844391	Hyla arenicolor Rag-1 gene, partial cds.		
DEFINITION	AY844391		
ACCESSION	AY844391.1		
VERSION	GI:61725398		
KEYWORDS	.		
SOURCE	Hyla arenicolor		
ORGANISM	Hyla arenicolor		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylinae; Hyla.	
REFERENCE	1 (bases 1 to 428)
AUTHORS	Faivovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A. and Wheeler, W.C.
TITLE	Systematic Review of the frog family Hylidae, with special reference to the Hylinae: Phylogenetic analysis and taxonomic revision
JOURNAL	Bull. Am. Mus. Nat. Hist. 294, 1-240 (2005)
REFERENCE	2 (bases 1 to 428)
AUTHORS	Faivovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A. and Wheeler, W.C.
TITLE	Direct Submission
JOURNAL	Submitted (01-DEC-2004) Herpetology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
FEATURES	Location/Qualifiers
source	1..428
	/organism="Hyla arenicolor"
	/mol_type="genomic DNA"
	/specimen_voucher="UMMZ 7755"
	/db_xref="taxon:251751"
	/country="USA: Arizona, Gila Co., Houston creek just N HWY 260, approx. 4 MI E of Payson"
mRNA	<1..>428
CDS	/product="Rag-1"
	<1..>428
	/codon_start=3
	/product="Rag-1"
	/protein_id="AA54101.1"
	/db_xref="GI:61725399"
	/translation="PVPERKKQATLDKLRKKNNLKPIMRMNGNFARKLMKSTVEA VCDLIPSEERQALREUMLDLYLKNKPVRRSSCPAKCEPCLLCQSYHSQRFAELLSTK FKRYEKTINYPFKTLHVPETIERDGSIGAWASEGNS"
ORIGIN	
Query Match	51.9%; Score 21.8; DB 11; Length 428;
Best Local Similarity	70.7%; Pred. No. 1.6e+02;
Matches	29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Qy	1 GCCACGCTGGCGCGCCCTTACAAAGACTATGAAGTAA 41
Db	29 GGCAACGCTTGACAAACACCTTAGAAGAGAGATGAAGTAA 69
RESULT 100	
AY844420	
LOCUS	AY844420 428 bp DNA linear VRT 26-AUG-2005
DEFINITION	Hyla japonica Rag-1 gene, partial cds.
ACCESSION	AY844420
VERSION	AY844420.1 GI:61725456
KEYWORDS	
SOURCE	Hyla japonica (Japanese treefrog)
ORGANISM	Hyla japonica
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylinae; Hyla.
REFERENCE	1 (bases 1 to 428)
AUTHORS	Faivovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A. and Wheeler, W.C.
TITLE	Systematic Review of the frog family Hylidae, with special reference to the Hylinae: Phylogenetic analysis and taxonomic revision
JOURNAL	Bull. Am. Mus. Nat. Hist. 294, 1-240 (2005)
REFERENCE	2 (bases 1 to 428)
AUTHORS	Faivovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A. and Wheeler, W.C.
TITLE	Direct Submission
JOURNAL	Submitted (01-DEC-2004) Herpetology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
FEATURES	Location/Qualifiers
source	1..428
	/organism="Hyla japonica"

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:16:40 ; Search time 312.2 Seconds

(without alignments)
937.971 Million cell updates/sec

Title: US-10-600-816-30

Perfect score: 42

Sequence: 1 gccacgcttgccgagccc.....acaaagactatgaagtaaaag 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10499840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : N Geneseq 8:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

14: geneseqn2005s:*

15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	42	100.0	497 11	ADT96418 Colon can
2	42	100.0	497 11	ADT96418 Colon can
3	42	100.0	552 11	ADT96292 Colon can
C 3	42	100.0	552 11	ADT96292 Colon can
C 4	42	100.0	552 11	ADT96292 Colon can
5	42	100.0	603 3	AAZ90046 Human col
C 6	42	100.0	620 6	ABQ58527 Human col
C 7	42	100.0	634 6	ABQ58527 Human col
C 8	42	100.0	642 10	ABT22328 Breast ca
9	42	100.0	1074 12	ADO30035 Human GPC
10	42	100.0	1212 3	AAZ61776 cDNA enco
11	42	100.0	1212 4	AAZ99709 Skin cell
12	42	100.0	1212 6	ABL34861 Human col
13	42	100.0	1228 8	ACA04775 cDNA enco
14	42	100.0	1460 14	AEA00112 Human TAT
15	42	100.0	1460 14	AEA00632 Human TAT
16	42	100.0	1619 4	AAF58615 Human REC
17	42	100.0	1718 3	AAZ90056 Hydrophob
18	42	100.0	2302 6	ABT10173 Human bre

Acc58386 Human GPC	2302	8	ACC58386
Adg93240 RAIGI cod	2302	10	ADD93240
Adl24773 Intestina	2302	10	ADL24773
Abt31923 Human bre	2316	10	ABT31923
Aah14688 Human GPC	2446	4	AAH14688
Abz42832 Human GPC	2456	8	ABZ42832
Acc72695 Human can	2456	10	ACC72695
Adn39301 Cancer/an	2456	11	ADN39301
Adl28525 Human GPC	2456	12	ADL28525
Adl28459 Human GPC	2456	12	ADL28459
Adl48221 Human ret	2456	13	ADL48221
Acn39388 Tumour-as	2456	13	ACN39388
Adl43992 Human bre	2456	13	ADL43992
Adu06126 Novel bro	2456	13	ADU06126
Aea00088 Human TAT	2456	14	AEA00088
Aea00608 Human TAT	2456	14	AEA00608
Aed47480 Retinoic	2456	14	AED47480
Abq54954 Human ova	2593	6	ABQ54954
Aah72766 Human cer	3371	4	AAH72766
Acn89274 Breast ca	4239	11	ACN89274
Adx97494 Pancreat	6730	13	ADX97494
Adf70573 Orphan re	1788	10	ADF70573
Aah08354 Human CDN	680	4	AAH08354
Adl28530 Human GPC	42	12	ADL28530
Adl28527 Human GPC	40	12	ADL28527
Adv92548 Japanese	480	14	ADV92548
Adv92553 Japanese	525	14	ADV92553
Aaa29595 Japanese	1500	3	AAA29595
Adv92550 Japanese	1500	14	ADV92550
Adv92562 Japanese	1545	14	ADV92562
Adm97133 Japanese	2088	12	ADM97133
Aaq22767 JEV Nakay	4512	2	AAQ22767
Aan80297 Partial n	4992	1	AAN80297
Adw74114 PCJEME DN	7486	14	ADW74114
Abk62169 Rat seque	584	6	ABK62169
Abd55155 Toxicity	52.9	584	10
Abd49676 Primary r	584	10	ABD49676
Adc85956 Human GPC	1339	10	ADC85956
Abx63126 Human CDN	4826	8	ABX63126
Adj56498 Human CDN	4826	10	ADJ56498
Abk42428 Genomic s	17849	4	ABK42428
Aal07451 Human rep	17849	4	AAL07451
Abd60584 Connectiv	17849	9	ABD60584
Abk42426 Genomic s	17862	4	ABK42426
Aal07450 Human rep	17862	4	AAL07450
Abd60582 Connectiv	17862	9	ABD60582
Abi17076 Drosophil	21748	4	ABI17076
Aed85840 Human CD3	363	14	AED85840
Aed85870 Human CD3	363	14	AED85870
Aac16864 Human sec	452	3	AAC16864
Aad42934 Human pho	70000	6	AAD42934
Aac28686 Human sec	407	3	AAC28686
Abi03149 Drosophil	2356	4	ABI03149
Adw52673 Japanese	882	14	ADW52673
Aas92414 DNA encod	1446	5	AAS92414
Adse0894 Bacterial	1555	13	ADS0894
Adi70879 Human ova	527	5	ADI70879
Adi77193 Human ova	527	5	ADI77193
Adi73027 Human ova	559	5	ADI73027
Adl38162 Human ova	559	5	ADL38162
Adl36609 Human ova	579	5	ADL36609
Adi71451 Human ova	579	5	ADI71451
Adl44546 Human ova	625	5	ADL44546
Adl44897 Human ova	692	5	ADL44897
Ado30325 Mouse GPC	1068	12	ADO30325
Adl62421 Human ova	1096	5	ADL62421
Adl62598 Human ova	1386	5	ADL62598
Adsl0225 Human the	1566	13	ADSL0225
Acn38091 Tumour-as	1580	13	ACN38091
Abn59642 Novel hum	1648	6	ABN59642
Adl62578 Human ova	2145	5	ADL62578
Adel11553 Human the	3870	13	ADE11553
Adi22610 Human liv	4605	10	ADI22610

92	21	50.0	6346	12	ADQ22756	Adq22756 Human sof	165	20	47.6	667	13	ADR62740	Adr62740 Cotton cd
93	21	50.0	7174	10	ADC37602	Adc37602 Human nuc	c 166	20	47.6	1123	6	ABL89772	AbL89772 Human pol
c 94	21	50.0	96589	9	ADA02708	Ada02708 Human ZFH	c 167	20	47.6	7690	6	ABL33123	AbL33123 Human imm
c 95	21	50.0	96589	10	ADB72446	AdB72446 Human ZFH	c 168	20	47.6	10591	12	ADO59810	AdO59810 Novel hum
c 96	21	50.0	96589	10	ADe95956	AdE95956 Human ZFH	c 169	20	47.6	38703	14	AE83346	Ae83346 Human cos
c 97	20.8	49.5	11728	10	ADC93459	Adc93459 E. faeciu	c 170	19.8	47.1	546	3	AAFO8124	AaFO8124 Fusarium cos
c 98	20.6	49.0	377	2	ADT23493	Adt23493 Human gen	c 171	19.8	47.1	546	13	ADU52165	AdU52165 Fusarium
c 99	20.6	49.0	457	9	ACH22183	Ach22183 Human adu	c 172	19.8	47.1	546	14	ADZ90168	AdZ90168 Fusarium
c 100	20.6	49.0	682	5	ADH45933	Adh45933 Human ova	c 173	19.8	47.1	726	6	ABN76605	AbN76605 Human syn
c 101	20.6	49.0	934	13	ADX60641	Adx60641 Plant ful	c 174	19.8	47.1	1125	13	ADT45575	AdT45575 Bacterial
c 102	20.6	49.0	963	4	AAE33117	Aae33117 Human sec	c 175	19.8	47.1	1128	13	ADS46516	AdS46516 Bacterial
c 103	20.6	49.0	1764	3	AAE77933	Aae77933 Human can	c 176	19.8	47.1	1405	6	AAK99775	AaK99775 1405at DN
c 104	20.6	49.0	2085	5	ABV28029	Abv28029 Human pro	c 177	19.8	47.1	1545	4	AAAF90506	AaAF90506
c 105	20.6	49.0	2085	5	ABV22192	Abv22192 Human pro	c 178	19.8	47.1	1545	10	ADB69816	AdB69816 C. neofor
c 106	20.6	49.0	2085	5	ABV27935	Abv27935 Human pro	c 179	19.8	47.1	1720	8	ABX63106	AbX63106 Human CDN
c 107	20.6	49.0	2085	5	ABV22095	Abv22095 Human pro	c 180	19.8	47.1	6683	10	ADB53086	AdB53086 Primary r
c 108	20.6	49.0	2387	13	ACN40823	Acn40823 Tumour-as	c 181	19.8	47.1	6683	13	ADV41291	AdV41291 Rat cardi
c 109	20.6	49.0	2610	3	AAZ90378	Aaz90378 Human dea	c 182	19.8	47.1	6874	4	AAAS45441	AaAS45441 Chemicall
c 110	20.6	49.0	2759	13	ACN40821	Acn40821 Tumour-as	c 183	19.8	47.1	6874	6	ABL70382	AbL70382 Chemicall
c 111	20.6	49.0	2759	14	ABE22826	AbE22826 Human col	c 184	19.8	47.1	6874	6	ABK28290	AbK28290 DNA trans
c 112	20.6	49.0	2772	4	AAH15102	AaH15102 Human CDN	c 185	19.8	47.1	6874	6	AAAS61334	AaAS61334 Human gen
c 113	20.6	49.0	2895	13	ADU01853	AdU01853 Novel hum	c 186	19.8	47.1	6874	6	ABN80231	AbN80231 Human che
c 114	20.6	49.0	2977	13	ACN40822	Acn40822 Tumour-as	c 187	19.8	47.1	10699	4	ADL14610	AdL14610 Wild-type
c 115	20.6	49.0	2977	14	ABE22793	AbE22793 Human col	c 188	19.8	47.1	10699	4	ADL14611	AdL14611 Attenuate
c 116	20.6	49.0	4279	14	ADX07488	AdX07488 Cyclin-de	c 189	19.8	47.1	11660	6	ABN85368	AbN85368 Spumaret
c 117	20.6	49.0	4506	15	ABE87869	Abe87869 Human CDN	c 190	19.8	47.1	143899	6	AAJ38336	AaJ38336 Genomic s
c 118	20.6	49.0	4522	5	AAAS64864	AaAS64864 DNA encod	c 191	19.8	46.7	49	14	ADV68377	AdV68377 Hevea bra
c 119	20.6	49.0	6579	8	ABT18965	Abt18965 Aspergill	c 192	19.6	46.7	272	12	ADQ06087	AdQ06087 Soybean t
c 120	20.6	49.0	7290	4	AAK53352	AaK53352 Human pol	c 193	19.6	46.7	337	8	AAV78634	AaV78634 Staphyloc
c 121	20.6	49.0	7838	4	AAK52368	AaK52368 Human pol	c 194	19.6	46.7	405	8	ACF73285	AcF73285 Staphyloc
c 122	20.6	49.0	30826	4	AAK67051	AaK67051 Human imm	c 195	19.6	46.7	539	10	ADD34794	AdD34794 Mouse nit
c 123	20.4	48.6	264	5	AAE68875	Aae68875 Human lun	c 196	19.6	46.7	692	13	ADX36505	AdX36505 Plant ful
c 124	20.4	48.6	264	6	ABK38786	Abk38786 cDNA clon	c 197	19.6	46.7	750	8	ACF64343	AcF64343 Human IFN
c 125	20.4	48.6	264	8	ACA11115	AcA11115 Human lun	c 198	19.6	46.7	879	10	ABS51668	AbS51668 LEC-1-rel
c 126	20.4	48.6	264	8	ACAO2301	AcAO2301 Lung canc	c 199	19.6	46.7	879	10	ADC23662	AdC23662 cDNA enco
c 127	20.4	48.6	264	10	ADH46342	AdH46342 Human lun	c 200	19.6	46.7	880	8	ACC00774	Acc00774 Glycine m
c 128	20.4	48.6	264	13	ADJ20261	AdJ20261 Human lun	c 201	19.6	46.7	946	13	ADX27655	AdX27655 Plant ful
c 129	20.4	48.6	363	14	ACL67473	AcL67473 M. xanthu	c 202	19.6	46.7	965	8	ACD20405	AcD20405 DNA encod
c 130	20.4	48.6	629	5	ABV55696	Abv55696 Human pro	c 203	19.6	46.7	988	2	AAQ70155	AaQ70155 Sequence
c 131	20.4	48.6	707	6	ABK78003	Abk78003 Bacillus	c 204	19.6	46.7	993	12	ADI42284	AdI42284 Plant tra
c 132	20.4	48.6	708	6	ABK77843	Abk77843 Bacillus	c 205	19.6	46.7	993	12	ADO02743	AdO02743 Soybean o
c 133	20.4	48.6	728	4	AAH06438	AaH06438 Human CDN	c 206	19.6	46.7	1020	3	ADO62332	AdO62332 Transcript
c 134	20.4	48.6	1659	13	ADT87784	Adt87784 Plant CDN	c 207	19.6	46.7	1020	3	AAA49789	Aaa49789 Arabidops
c 135	20.4	48.6	2318	4	AAK53300	AaK53300 Human pol	c 208	19.6	46.7	1222	2	AAQ70158	AaQ70158 Arabidops
c 136	20.4	48.6	2523	10	ADA53709	Ada53709 Human cod	c 209	19.6	46.7	1344	8	ACA37293	AcA37293 Prokaryot
c 137	20.4	48.6	2623	4	AAK52316	AaK52316 Human pol	c 210	19.6	46.7	1428	10	ABT42343	AbT42343 Toxicity
c 138	20.4	48.6	2648	10	ADB62195	AdB62195 Human CDN	c 211	19.6	46.7	1428	13	ADV41584	AdV41584 Rat cardi
c 139	20.4	48.6	2737	4	ABL24858	AbL24858 Drosophi	c 212	19.6	46.7	1500	1	AAW71394	AaW71394 Flaviviru
c 140	20.4	48.6	3057	15	AAE88935	Aae88935 Human che	c 213	19.6	46.7	1531	2	AAQ70156	AaQ70156 Sequence
c 141	20.4	48.6	3240	4	AAH14453	AaH14453 Human CDN	c 214	19.6	46.7	2040	4	AAH26533	AaH26533 Human pro
c 142	20.4	48.6	21295	14	ACL64731	AcL64731 M. xanthu	c 215	19.6	46.7	2054	13	ADW23987	AdW23987 Japanese
c 143	20.4	48.6	122614	11	ACN44998	Acn44998 Human gen	c 216	19.6	46.7	2070	4	AAH26532	AaH26532 Human pro
c 144	20.4	48.6	167932	10	ADL13501	AdL13501 Osteoarth	c 217	19.6	46.7	2771	5	ABA18382	AbA18382 Human ner
c 145	20.4	48.6	265118	5	AAH41227	AaH41227 Pyrococcu	c 218	19.6	46.7	3110	13	ADW23445	AdW23445 Japanese
c 146	20.2	48.1	384	6	ABE54351	AbE54351 DNA encod	c 219	19.6	46.7	3577	6	ABT199459	AbT199459 Mouse isc
c 147	20.2	48.1	614	3	AAAC45316	AaAC45316 Arabidops	c 220	19.6	46.7	5293	10	ABV77549	AbV77549 Plasmid p
c 148	20.2	48.1	1410	8	ACA41716	AcA41716 Prokaryot	c 221	19.6	46.7	5334	10	ABV77537	AbV77537 Plasmid p
c 149	20.2	48.1	1419	3	AAAC44553	AaAC44553 Arabidops	c 222	19.6	46.7	7500	10	ABV77533	AbV77533 Japanese
c 150	20.2	48.1	1545	6	AAAS97238	AaAS97238 Neisseria	c 223	19.6	46.7	9310	2	AAV74627	AaV74627 Staphyloc
c 151	20.2	48.1	1548	3	AA254374	Aa254374 Neisseria	c 224	19.6	46.7	10892	4	AAAF83821	AaAF83821 Chimeric
c 152	20.2	48.1	5587	10	ADAE5229	AdAE5229 Rat gene	c 225	19.6	46.7	10892	14	ABE66292	AbE66292 Yellow fe
c 153	20.2	48.1	11851	12	ADL98344	AdL98344 lin(in)3628	c 226	19.6	46.7	10976	3	ABL50890	AbL50890 Japanese
c 154	20.2	48.1	108359	9	ADA13316	AdA13316 Human fri	c 227	19.6	46.7	10985	4	AAK86420	AaK86420 Human imm
c 155	20	47.6	489	13	ACN57177	AcN57177 Cotton gy	c 228	19.6	46.7	10985	4	AAK86422	AaK86422 Human imm
c 156	20	47.6	527	4	AAI16854	AaI16854 Human bre	c 229	19.6	46.7	10985	4	AAK86421	AaK86421 Human imm
c 157	20	47.6	531	6	AAK98730	AaK98730 DNA seque	c 230	19.6	46.7	41038	14	ABE96523	AbE96523 Human IFN
c 158	20	47.6	542	4	AAO03843	AaO03843 Complemen	c 231	19.6	46.7	51558	13	ACN37207	AcN37207 Human per
c 159	20	47.6	542	4	AAO03842	AaO03842 FLAG-huma	c 232	19.6	46.7	52677	11	ACN43934	AcN43934 Human gen
c 160	20	47.6	546	6	AAK98733	AaK98733 DNA of a	c 233	19.6	46.7	87320	14	ADZ12863	AdZ12863 Human can
c 161	20	47.6	576	2	AAQ83893	AaQ83893 Hepatitis	c 234	19.6	46.7	110000	14	ABE39172	AbE39172 Continuation (2 of
c 162	20	47.6	576	2	AAAT16607	AaAT16607 Hepatitis	c 235	19.6	46.7	110000	14	ABE42736	AbE42736 Continuation (5 of
c 163	20	47.6	608	13	ACN61430	AcN61430 Cotton gy	c 236	19.6	46.7	349980	6	ABQ81849	AbQ81849 Bifidobac
c 164	20	47.6	649	5	ADL42993	AdL42993 Human ova	c 237	19.4	46.2	480	13	ACF87190	AcF87190 Human SIR

C 238	19.4	46.2	546	12	ADJ10916	Adi10916 Recombina	C 311	19.2	45.7	2325	8	ACC46631	Acc46631 Human dit
C 239	19.4	46.2	570	5	ADL33095	Adl33095 Human ova	C 312	19.2	45.7	2428	4	AAH98381	Aah98381 Human EST
C 240	19.4	46.2	810	6	ABL06697	Abi06697 Human dit	C 313	19.2	45.7	2428	4	AAH98396	Aah98396 Human EST
C 241	19.4	46.2	844	4	ABL06697	Abi06697 Human dit	C 314	19.2	45.7	2451	2	AAV15809	Aav15809 Micrococ
C 242	19.4	46.2	883	10	ADBS8936	Adbs8936 Toxicity-	C 315	19.2	45.7	3144	11	ADM03175	Adm03175 Human cDN
C 243	19.4	46.2	883	10	ABT41733	Abt41733 Toxicity	C 316	19.2	45.7	3144	11	AEC86105	Aec86105 Human cDN
C 244	19.4	46.2	883	12	ADP72564	Adp72564 Renal tox	C 317	19.2	45.7	3257	3	AAZ52447	Aaz52447 HTRM clon
C 245	19.4	46.2	1682	8	ACA22852	Acas22852 Prokaryot	C 318	19.2	45.7	3500	13	ADR07983	Adr07983 Full leng
C 246	19.4	46.2	1725	3	AAA96701	Aaa96701 Reporart	C 319	19.2	45.7	3507	14	ADX07941	Adx07941 Cyclin-de
C 247	19.4	46.2	2173	4	ABL06696	Abi06696 Drosophil	C 320	19.2	45.7	3507	14	ADZ49474	Adz49474 Insulin s
C 248	19.4	46.2	2425	5	AAS79360	Aas79360 DNA encod	C 321	19.2	45.7	3675	4	ABL14979	Abi14979 Drosophil
C 249	19.4	46.2	2499	10	ADF28508	Adf28508 NgrHy con	C 322	19.2	45.7	4005	8	ABT18026	Abt18026 Aspergill
C 250	19.4	46.2	2565	10	ADE06964	Ade06964 Novel cod	C 323	19.2	45.7	4235	8	ABT19840	Abt19840 Aspergill
C 251	19.4	46.2	3142	4	ABL14842	Abi14842 Drosophil	C 324	19.2	45.7	4300	4	AAH42107	Aah42107 Nucleotid
C 252	19.4	46.2	3750	15	AEE26295	Aee26295 BRCT doma	C 325	19.2	45.7	4309	14	AEA27507	Aea27507 Human zin
C 253	19.4	46.2	3895	11	ACN44011	Acn44011 Human mRN	C 326	19.2	45.7	4909	14	AEA30110	Aea30110 Glioblast
C 254	19.4	46.2	6878	4	ABL17536	Abi17536 Drosophil	C 327	19.2	45.7	4960	6	ABS73413	Abs73413 Human Gll
C 255	19.4	46.2	7430	4	AAK77418	Aak77418 Human imm	C 328	19.2	45.7	5191	10	ADH28900	Adh28900 Human chr
C 256	19.4	46.2	7894	8	ABZ76334	Abz76334 L. cuprin	C 329	19.2	45.7	5191	12	ADG39645	Adg39645 Human cDN
C 257	19.4	46.2	10002	8	ABZ76333	Abz76333 L. cuprin	C 330	19.2	45.7	5191	12	ADG39694	Adg39694 Human cDN
C 258	19.4	46.2	10341	8	ABZ76340	Abz76340 L. cuprin	C 331	19.2	45.7	5264	6	AAZ37432	Aaz37432 PCDD2/act
C 259	19.4	46.2	20210	4	AAK77417	Aak77417 Human imm	C 332	19.2	45.7	5490	6	ABQ99289	Abq99289 Human cod
C 260	19.4	46.2	20210	4	AAS31519	Aas31519 Human DNA	C 333	19.2	45.7	5906	15	AEF74614	Aef74614 Human pol
C 261	19.4	46.2	20210	6	ABQ66843	Abq66843 Human pol	C 334	19.2	45.7	5981	8	ABX63291	Abx63291 Human cDN
C 262	19.4	46.2	20210	10	ADC11130	Adc11130 Human DNA	C 335	19.2	45.7	7619	4	ABL14978	Abi14978 Drosophil
C 263	19.4	46.2	38019	14	AES04707	Aee04707 Cancer-as	C 336	19.2	45.7	9027	8	AAH49171	Aah49171 CFT1030 v
C 264	19.4	46.2	83836	11	ACN45080	Acn45080 Mouse gen	C 337	19.2	45.7	9259	8	AAH49174	Aah49174 CFT1030 v
C 265	19.4	46.2	108526	14	AD213294	Ad213294 Murine ca	C 338	19.2	45.7	10553	8	AAH49160	Aah49160 CFT710 ve
C 266	19.4	46.2	110000	14	AEA61120_4	Continuation (5 of	C 339	19.2	45.7	38358	4	AAK73535	Aak73535 Human imm
C 267	19.4	46.2	110000	14	AEA61120_2	Continuation (3 of	C 340	19.2	45.7	106323	8	ABX14591	Abx14591 Human chl
C 268	19.4	46.2	119854	14	AD213282	Ad213282 Murine ca	C 341	19.2	45.7	106323	10	AAD58821	Aad58821 Human tra
C 269	19.4	46.2	157090	12	AD047194	Ad047194 DNA seque	C 342	19.2	45.7	107159	12	ADL13431	Adl13431 Human thy
C 270	19.4	46.2	196686	11	ACN44170	Acn44170 Human gen	C 343	19.2	45.7	110000	10	ACF67367_22	Continuation (23 o
C 271	19.4	46.2	198522	11	ACN44010	Acn44010 Human gen	C 344	19.2	45.7	110000	10	ACF67367_23	Continuation (24 o
C 272	19.4	46.2	325093	14	AEE25355	Aee25355 Human ARR	C 345	19.2	45.7	110000	10	ACF65386_5	Continuation (6 of
C 273	19.2	45.7	252	5	ABV07221	Abv07221 Human pro	C 346	19.2	45.7	139308	8	ADB12769	Adb12769 Human PRK
C 274	19.2	45.7	395	4	AAI84093	Aai84093 Human pol	C 347	19.2	45.7	155083	12	ADQ97891	Adq97891 Mouse can
C 275	19.2	45.7	398	5	ABV27161	Abv27161 Human pro	C 348	19.2	45.7	202251	11	ACN44504	Acn44504 Mouse gen
C 276	19.2	45.7	450	11	ABD09803	Abd09803 Pseudomon	C 349	19.2	45.7	202251	12	ADL13431	Adl13431 Human thy
C 277	19.2	45.7	459	10	ACF69385	Acf69385 Pseudorhab	C 350	19.2	45.7	289	6	ABL71178	Abi71178 Corn tass
C 278	19.2	45.7	465	4	AAI15537	Aai15537 Probe #54	C 351	19.2	45.7	297	8	ABZ22706	Abz22706 Human ant
C 279	19.2	45.7	465	4	ABA57440	Aba57440 Human foe	C 352	19.2	45.7	456	9	ACH40353	Ach40353 Human foe
C 280	19.2	45.7	465	4	AAI36988	Aai36988 Probe #56	C 353	19.2	45.7	457	13	ACF84030	Acf84030 Human STR
C 281	19.2	45.7	465	4	ABA26945	Aba26945 Probe #54	C 354	19.2	45.7	471	10	ADE59600	Ade59600 Rat gene
C 282	19.2	45.7	465	4	AAK31079	Aak31079 Human bon	C 355	19.2	45.7	482	9	ACH41862	Ach41862 Human foe
C 283	19.2	45.7	465	4	AAK05478	Aak05478 Human bra	C 356	19.2	45.7	489	5	AAH81814	Aah81814 Human dif
C 284	19.2	45.7	465	4	ABS30758	Abs30758 Human liv	C 357	19.2	45.7	489	5	AAH81638	Aah81638 Human dif
C 285	19.2	45.7	465	6	ABS05830	Abs05830 Human gen	C 358	19.2	45.7	514	8	ABZ53921	Abz53921 Aspergill
C 286	19.2	45.7	507	13	ADR63029	Adr63029 Cotton cD	C 359	19.2	45.7	529	13	ACF91105	Acf91105 Human STR
C 287	19.2	45.7	578	13	ACN49195	Acn49195 Cotton pr	C 360	19.2	45.7	536	13	ADQ54362	Adq54362 Novel can
C 288	19.2	45.7	579	2	AAQ78092	Aaq78092 Hepatitis	C 361	19.2	45.7	566	4	AAH13627	Aah13627 Human cDN
C 289	19.2	45.7	579	2	AAQ78113	Aaq78113 Hepatitis	C 362	19.2	45.7	576	2	AAQ83888	Aaq83888 Hepatitis
C 290	19.2	45.7	579	2	AAQ78093	Aaq78093 Hepatitis	C 363	19.2	45.7	576	2	AAH16602	Aat16602 Hepatitis
C 291	19.2	45.7	580	2	AAQ78045	Aaq78045 Hepatitis	C 364	19.2	45.7	1188	2	AAQ22974	Aaq22974 Sequence
C 292	19.2	45.7	580	2	AAQ78046	Aaq78046 Hepatitis	C 365	19.2	45.7	1224	13	ADR63445	Adr63445 Cotton cD
C 293	19.2	45.7	630	2	AAI12966	Aai12966 HCV El co	C 366	19.2	45.7	1260	2	AAQ22972	Aaq22972 Sequence
C 294	19.2	45.7	630	6	AAH48930	Aah48930 Hepatitis	C 367	19.2	45.7	1263	14	ADV42778	Adv42778 Human psy
C 295	19.2	45.7	630	10	ADD55539	Add55539 Hepatitis	C 368	19.2	45.7	1327	3	AAA34861	Aaa34861 Human low
C 296	19.2	45.7	630	12	ADP71121	Adp71121 HCV DNA e	C 369	19.2	45.7	1327	3	AAH20983	Aah20983 Human low
C 297	19.2	45.7	959	2	AAQ78049	Aaq78049 Hepatitis	C 370	19.2	45.7	1327	10	ABZ96677	Abz96677 Human nuc
C 298	19.2	45.7	959	2	AAQ78047	Aaq78047 Hepatitis	C 371	19.2	45.7	1327	11	ABD20526	Abd20526 Human pul
C 299	19.2	45.7	959	2	AAQ78048	Aaq78048 Hepatitis	C 372	19.2	45.7	1351	2	AAQ25789	Aaq25789 Encodes s
C 300	19.2	45.7	1311	13	AAQ63027	Aaq63027 Cotton cD	C 373	19.2	45.7	1351	2	AAQ38440	Aaq38440 SHL1-SR-a
C 301	19.2	45.7	1491	3	AACT76674	Aact76674 Human ORF	C 374	19.2	45.7	1372	11	AEA90313	Aea90313 Human NOV
C 302	19.2	45.7	1641	8	ABT19214	Abt19214 Aspergill	C 375	19.2	45.7	1405	3	AAA34864	Aaa34864 Human low
C 303	19.2	45.7	1683	11	ABD09733	Abd09733 Pseudomon	C 376	19.2	45.7	1405	3	AAQ20986	Aaq20986 Human low
C 304	19.2	45.7	1725	13	ADS14550	Ads14550 Pseudomon	C 377	19.2	45.7	1405	10	ABZ96680	Abz96680 Human nuc
C 305	19.2	45.7	1833	11	ABD09767	Abd09767 Pseudomon	C 378	19.2	45.7	1405	11	ABD20529	Abd20529 Human pul
C 306	19.2	45.7	1980	11	ABD09903	Abd09903 Pseudomon	C 379	19.2	45.7	1608	2	AAQ25790	Aaq25790 Encodes n
C 307	19.2	45.7	2000	11	ACL36039	Acl36039 Rice stre	C 380	19.2	45.7	1664	3	AAA34862	Aaa34862 Human low
C 308	19.2	45.7	2006	8	ABT18620	Abt18620 Aspergill	C 381	19.2	45.7	1664	3	AAQ20984	Aaq20984 Human low
C 309	19.2	45.7	2070	8	ABT21034	Abt21034 Aspergill	C 382	19.2	45.7	1664	10	ABZ96678	Abz96678 Human nuc
C 310	19.2	45.7	2235	8	ABT20436	Abt20436 Aspergill	C 383	19.2	45.7	1664	10	ACA56622	Acc56622 Human sig

384	19	45.2	1664	11	ABD20527	Abd20527 Human pul	c 457	18.8	44.8	412	6	ABN23726	Abn23726 Human ORF
385	19	45.2	1664	12	AD156418	Ad156418 Human pol	458	18.8	44.8	413	10	ADD26877	Add26877 Human adi
386	19	45.2	1784	14	ABE67638	Aeb67638 Rice geno	c 459	18.8	44.8	466	8	ABX41497	Abx41497 Bovine ES
387	19	45.2	1899	2	AAT05530	Aat05530 Metallopro	c 460	18.8	44.8	508	12	ACH77040	Ach77040 Human gen
388	19	45.2	1998	3	AAA34863	Aaa34863 Human ade	c 461	18.8	44.8	540	8	ACF72639	Acf72639 Staphyloc
389	19	45.2	1998	3	AAC73842	Aac73842 Human IL-	c 462	18.8	44.8	551	5	ABA12328	Abal2328 Human ner
390	19	45.2	1998	3	AAT20985	Aat20985 Human low	c 463	18.8	44.8	563	3	AAA42386	Aaa42386 Human sec
391	19	45.2	1998	8	ABX04496	Abx04496 Human int	c 464	18.8	44.8	597	4	AAI16648	Aai16648 Probe #65
392	19	45.2	1998	10	ABZ96679	Abz96679 Human nuc	c 465	18.8	44.8	597	4	ABA59987	Abas9987 Human foe
393	19	45.2	1998	11	ABD20528	Abd20528 Human pul	c 466	18.8	44.8	597	4	AAI139860	Aai139860 Probe #85
394	19	45.2	1998	12	ADR12173	Adr12173 Human int	c 467	18.8	44.8	597	4	ABA28396	Abas28396 Probe #68
395	19	45.2	2006	2	AAQ22978	Aaq22978 Sequence	c 468	18.8	44.8	597	4	AAK34137	Aak34137 Human bon
396	19	45.2	2006	2	AAQ22973	Aaq22973 Sequence	c 469	18.8	44.8	597	4	AAK08258	Aak08258 Human bra
397	19	45.2	2006	3	AAA34865	Aaa34865 Human ade	c 470	18.8	44.8	597	4	ABS33935	Abs33935 Human liv
398	19	45.2	2006	3	AAC73823	Aac73823 Human IL-	c 471	18.8	44.8	597	6	ABS08900	Abas08900 Human gen
399	19	45.2	2006	3	AAT20987	Aat20987 Human low	c 472	18.8	44.8	633	6	ABQ42805	Abq42805 Oligonuel
400	19	45.2	2006	8	ABX04477	Abx04477 Human int	c 473	18.8	44.8	633	6	ABQ42804	Abq42804 Oligonuel
401	19	45.2	2006	10	ABZ96681	Abz96681 Human nuc	c 474	18.8	44.8	698	4	AAH06169	Aah06169 Human CDN
402	19	45.2	2006	11	ABD20530	Abd20530 Human pul	c 475	18.8	44.8	750	2	AAV24854	Aav24854 H. pylori
403	19	45.2	2006	12	ADJ61308	Adj61308 ILSR-X611	c 476	18.8	44.8	750	2	AAK75831	Aak75831 H. pylori
404	19	45.2	2006	12	ADO46698	Ado46698 Human oli	c 477	18.8	44.8	757	4	AAK93303	Aak93303 Human CDN
405	19	45.2	2006	12	ADR12154	Adr12154 Human int	c 478	18.8	44.8	757	4	AAK91894	Aak91894 Human CDN
406	19	45.2	2024	2	AAQ22979	Aaq22979 Sequence	c 479	18.8	44.8	757	12	ADL28321	Adl28321 5' end of
407	19	45.2	2024	2	AAQ22975	Aaq22975 Sequence	c 480	18.8	44.8	757	12	ADL29730	Adl29730 5' end of
408	19	45.2	2024	3	AAA34859	Aaa34859 Human ade	c 481	18.8	44.8	763	4	AAH99124	Aah99124 Human EST
409	19	45.2	2024	3	AAF20981	Aaf20981 Human low	c 482	18.8	44.8	780	2	AAV24741	Aav24741 H. pylori
410	19	45.2	2024	10	ABZ96675	Abz96675 Human nuc	c 483	18.8	44.8	1000	14	ABE85694	Aeb85694 Human pho
411	19	45.2	2024	11	ADI31989	Adi31989 Human CDN	c 484	18.8	44.8	1000	14	ABE85696	Aeb85696 Human pho
412	19	45.2	2024	11	ABD20524	Abd20524 Human pul	c 485	18.8	44.8	1029	9	ACF25382	Acf25382 Rat schwa
413	19	45.2	2024	12	ADJ59303	Adj59303 Interleuk	c 486	18.8	44.8	1251	3	AAK81855	Aax81855 cDNA enco
414	19	45.2	2024	12	ADO4793	Ado4793 Human oli	c 487	18.8	44.8	1251	3	AAA08002	Aaa08002 Human Ste
415	19	45.2	2024	13	ADS84056	AdS84056 Human lym	c 488	18.8	44.8	1289	6	ABO52169	AbO52169 Oligonuel
416	19	45.2	2052	2	AAT05529	Aat05529 p45 metal	c 489	18.8	44.8	1289	6	ABO52168	AbO52168 Oligonuel
417	19	45.2	2052	2	AAT40133	Aat40133 Fusarium	c 490	18.8	44.8	1346	9	ADA11561	Ada11561 Human CDN
418	19	45.2	2052	2	AAV28868	Aav28868 Fumarium	c 491	18.8	44.8	1346	9	ADA11561	Ada11561 Human CDN
419	19	45.2	2052	4	ABL04747	AbL04747 Drosophil	c 492	18.8	44.8	1353	3	AAK81858	Aax81858 Ampified
420	19	45.2	2212	8	AAU53176	Aau53176 Human kin	c 493	18.8	44.8	1353	3	AAA08005	Aaa08005 Human Ste
421	19	45.2	2322	6	ABA92185	AbA92185 Human NAL	c 494	18.8	44.8	1371	10	ADE56431	AdE56431 Rat gene
422	19	45.2	2322	6	ABA92206	AbA92206 Human NYX	c 495	18.8	44.8	1372	10	ADF82068	AdF82068 Leukaemia
423	19	45.2	2397	10	ADB62505	AdB62505 Human CDN	c 496	18.8	44.8	1429	10	ADK56649	AdK56649 Plant DNA
424	19	45.2	2547	5	AA575388	Aa575388 DNA enco	c 497	18.8	44.8	1495	3	ACA46957	Aac46957 Arabidops
425	19	45.2	2713	10	ADP69147	AdP69147 Human MP5	c 498	18.8	44.8	1737	5	AAF29901	Aaf29901 Human h22
426	19	45.2	2876	11	ADM02841	Adm02841 Human CDN	c 499	18.8	44.8	1770	4	AAH21020	Aah21020 Bovine-de
427	19	45.2	2876	14	AEC85771	Aec85771 Human CDN	c 500	18.8	44.8	1770	13	ADT15581	Adt15581 Plant CDN
428	19	45.2	3737	2	AAV43796	Aav43796 Sequence	c 501	18.8	44.8	2000	8	ADA72758	Ada72758 Rice gene
429	19	45.2	4345	4	ABL04746	AbL04746 Staphyloc	c 502	18.8	44.8	2436	8	ABT19144	Abt19144 Aspergill
430	19	45.2	5648	4	AAK89353	Aak89353 Staphyloc	c 503	18.8	44.8	2552	4	AAH16154	Aah16154 Human CDN
431	19	45.2	11355	3	AAA34866	Aaa34866 Human ade	c 504	18.8	44.8	2589	4	AAK94285	Aak94285 Human ful
432	19	45.2	11355	3	AAAT20988	Aat20988 Human low	c 505	18.8	44.8	2589	12	ADL30895	Adl30895 Full leng
433	19	45.2	11355	10	ABZ96682	Abz96682 Human nuc	c 506	18.8	44.8	2727	8	ABT18550	Abt18550 Aspergill
434	19	45.2	11355	11	ABD20531	Abd20531 Human pul	c 507	18.8	44.8	2800	14	AED04233	Aed04233 Human HGI
435	19	45.2	28140	12	ADO47032	Ado47032 Human oli	c 508	18.8	44.8	3020	4	AAK94468	Aak94468 Human ful
436	19	45.2	39700	12	AQD97170	Aqd97170 Mouse can	c 509	18.8	44.8	3020	12	ADL31251	Adl31251 Full leng
437	19	45.2	42587	12	ADJ61642	Adj61642 Concateme	c 510	18.8	44.8	3075	14	AED73704	Aed73704 Human pla
438	19	45.2	101169	12	AQD97584	Aqd97584 Mouse can	c 511	18.8	44.8	3160	8	ACC50232	Acc50232 Breast ca
439	19	45.2	110000	14	ADN76257	AdN76257 Continuation (4 of	c 512	18.8	44.8	3160	13	ADR25792	Adr25792 Breast ca
440	19	45.2	135090	11	ADN95683	AdN95683 Human BEC	c 513	18.8	44.8	3179	4	AAK94508	Aak94508 Human tum
441	19	45.2	210920	12	AQD97123	Aqd97123 Mouse can	c 514	18.8	44.8	3179	10	ADD29652	Add29652 Human ful
442	18.8	44.8	80	12	ADM95112	Adm95112 Rat antis	c 515	18.8	44.8	3179	12	ADL31329	Adl31329 Full leng
443	18.8	44.8	131	4	AAI25844	Aai25844 Probe #15	c 516	18.8	44.8	3200	2	AAK81854	Aax81854 cDNA enco
444	18.8	44.8	131	4	ABA72528	AbA72528 Human foe	c 517	18.8	44.8	3201	3	AAA08001	Aaa08001 Human Ste
445	18.8	44.8	131	4	AAI52941	Aai52941 Probe #21	c 518	18.8	44.8	3213	8	ABT20964	Abt20964 Aspergill
446	18.8	44.8	131	4	ABA38275	AbA38275 Probe #16	c 519	18.8	44.8	3263	10	ADC49713	Adc49713 Human WAS
447	18.8	44.8	131	4	AAK47105	Aak47105 Human bon	c 520	18.8	44.8	3263	13	ADP24956	Adp24956 PRO polyp
448	18.8	44.8	131	4	AAK20952	Aak20952 Human bra	c 521	18.8	44.8	3263	13	ADU05844	Adu05844 Novel bro
449	18.8	44.8	131	4	ABS46865	Abs46865 Human liv	c 522	18.8	44.8	3263	14	ADZ10045	Adz10045 Human bre
450	18.8	44.8	131	6	ABS21334	Abs21334 Human gen	c 523	18.8	44.8	3263	15	ABE87871	Aee87871 Human CDN
451	18.8	44.8	239	4	AAH27148	Aah27148 Human lep	c 524	18.8	44.8	3268	2	AAZ40479	Aaz40479 Human STL
452	18.8	44.8	243	6	ABN25451	Abn25451 Human ORF	c 525	18.8	44.8	3283	4	AAF75337	Aaf75337 Human TGF
453	18.8	44.8	279	12	ACH90757	Ach90757 Human gen	c 526	18.8	44.8	3335	10	ABE31758	AbE31758 Human 225
454	18.8	44.8	288	14	AEC38392	Aec38392 DNA enco	c 527	18.8	44.8	3363	8	ABT20366	Abt20366 Aspergill
455	18.8	44.8	288	14	AEC38390	Aec38390 DNA enco	c 528	18.8	44.8	3403	14	ABE32284	Aeb32284 Human CDN
456	18.8	44.8	393	6	ADH31558	Adh31558 Novel yea	c 529	18.8	44.8	3408	4	AAD17487	Aad17487 Human lep

C 530	18.8	44.8	3426	2	AAX59851	Aax59851 SQ ID 4	603	18.6	44.3	2271	4	ABL21710	Ab121710 Drosophil
C 531	18.8	44.8	3426	6	ABL63717	Ab163717 Breast ca	C 604	18.6	44.3	2396	4	ABL30488	Ab130488 Drosophil
C 532	18.8	44.8	3426	12	AD019662	Ad019662 Human sof	C 605	18.6	44.3	2396	4	ABL30488	Ab130488 Drosophil
C 533	18.8	44.8	3426	14	AEB32239	Aeb32239 Human cDN	C 606	18.6	44.3	2604	13	ADS58014	Ad58014 Bacterial
C 534	18.8	44.8	3560	4	AAL159504	Aal159504 Human act	C 607	18.6	44.3	2615	4	ABL09765	Ab109765 Drosophil
C 535	18.8	44.8	3755	8	ABX62955	Abx62955 Human pol	C 608	18.6	44.3	3054	4	ABL11272	Ab111272 Drosophil
C 536	18.8	44.8	4130	12	AD077047	Ad077047 Human cDN	C 609	18.6	44.3	3306	4	ABL13840	Ab113840 Drosophil
C 537	18.8	44.8	4258	5	AAS78055	Aas78055 DNA encod	C 610	18.6	44.3	3411	13	ADR07076	Adr07076 Pull leng
C 538	18.8	44.8	4622	4	ABK42254	Abk42254 Genomic s	C 611	18.6	44.3	3478	14	AEB26903	Aeb26903 Pinus rad
C 539	18.8	44.8	4622	9	ADB60410	Adb60410 Connectiv s	C 612	18.6	44.3	4570	4	AFF88172	Aff88172 A thalia
C 540	18.8	44.8	4727	8	ABT117956	Abt117956 Aspergill	C 613	18.6	44.3	4806	4	ABL09764	Ab109764 Drosophil
C 541	18.8	44.8	5363	8	ABT119770	Abt119770 Aspergill	C 614	18.6	44.3	5082	4	ABL04626	Ab104626 Drosophil
C 542	18.8	44.8	6240	4	ABL06443	Ab106443 Drosophil	C 615	18.6	44.3	6975	4	ABL09610	Ab109610 Drosophil
C 543	18.8	44.8	9644	14	AEB32406	Aeb32406 Human gen	C 616	18.6	44.3	10343	4	ABL14246	Ab114246 Drosophil
C 544	18.8	44.8	9933	4	ABL06442	Ab106442 Drosophil	C 617	18.6	44.3	11006	4	AAK66281	Aak66281 Human imm
C 545	18.8	44.8	10489	4	ABL14889	Ab114889 Drosophil	C 618	18.6	44.3	24174	14	AEE04893	Aee04893 Cancer-as
C 546	18.8	44.8	11566	5	ABAI4380	Abai4380 Human her	C 619	18.6	44.3	28136	4	AAK69755	Aak69755 Human imm
C 547	18.8	44.8	13484	4	ABLI14888	Abli14888 Drosophil	C 620	18.6	44.3	53267	14	ABE05141	AbE05141
C 548	18.8	44.8	13502	5	AAH48508	Aah48508 Japanese	C 621	18.6	44.3	61510	14	ADZ13267	Adz13267 Human can
C 549	18.8	44.8	20348	14	AEB32361	Aeb32361 Human gen	C 622	18.6	44.3	82256	12	ADO34927	Ado34927 3
C 550	18.8	44.8	39746	4	ABL13398	Ab113398 Drosophil	C 623	18.6	44.3	96583	3	AAF22297	Aaf22297 BAC conta
C 551	18.8	44.8	64125	13	ABD33295	Abd33295 Human can	C 624	18.6	44.3	96593	11	ADL27146	Adl27146 Human gen
C 552	18.8	44.8	106416	4	ABLI18718	Abli18718 Drosophil	C 625	18.6	44.3	96595	9	ADA03068	Ada03068 Human PPP
C 553	18.8	44.8	110000	2	AAX20248_05	Aax20248_05	C 626	18.6	44.3	96595	10	ADB72806	Adb72806 Human PPP
C 554	18.8	44.8	110000	2	AAX20248_06	Aax20248_06	C 627	18.6	44.3	96596	9	ADA66352	Ada66352 Human PPP
C 555	18.8	44.8	110000	6	ABX08336_01	Abx08336_01	C 628	18.6	44.3	110000	13	ABD32804	Abd32804 Mouse can
C 556	18.8	44.8	110000	12	ADJ25985_01	Adj25985_01	C 629	18.6	44.3	127168	15	ABF75295	Abf75295 Human pol
C 557	18.8	44.8	110000	12	ADN97989_01	Adn97989_01	C 630	18.6	44.3	127369	11	ACN44006	Acn44006 Human gen
C 558	18.8	44.8	110000	13	AD050281_01	Ad050281_01	C 631	18.6	44.3	199719	14	ABE05001	Abe05001 Cancer-as
C 559	18.8	44.8	110000	13	ABD32804_2	Abd32804_2	C 632	18.6	44.3	215221	11	ACN44754	Acn44754 Human gen
C 560	18.8	44.8	111309	14	AEB85185_01	Aeb85185_01	C 633	18.6	44.3	218336	8	ABQ76678	Abq76678 Androgen
C 561	18.8	44.8	111309	2	AAX20250	Aax20250 Borrelia	C 634	18.6	44.3	236303	4	AAS11614	Aas11614 Human gen
C 562	18.8	44.8	127222	12	ADQ97301	Adq97301 Human can	C 635	18.6	44.3	310122	13	ABD32533	Abd32533 Mouse can
C 563	18.8	44.8	152501	12	ADP677269	Adp677269 Human chr	C 636	18.6	44.3	310122	14	ADZ13032	Adz13032 Murine ca
C 564	18.8	44.8	156296	14	AEE05181	Aee05181 Cancer-as	C 637	18.6	44.3	310122	14	ADZ13032	Adz13032 Murine ca
C 565	18.8	44.8	160274	14	AEB32377	Aeb32377 Human gen	C 638	18.6	44.3	3227	2	AAK63450	Aak63450 Human imm
C 566	18.8	44.8	160300	14	AEB32388	Aeb32388 Human gen	C 639	18.6	44.3	3300	2	AAZ14924	Aaz14924 Human gen
C 567	18.8	44.8	177851	8	AAL57272	Aal57272 ba438B23-	C 640	18.6	44.3	3397	4	AAI82129	Aai82129 Human pol
C 568	18.8	44.8	262090	12	ADQ59207	Adq59207 MSI-H car	C 641	18.6	44.3	418	14	ACL54457	ACL54457 Human col
C 569	18.8	44.8	302603	11	ADP75187	Adp75187 Human end	C 642	18.6	44.3	436	10	ABD50391	Abd50391 Primary r
C 570	18.6	44.3	23	12	AD128532	Adi28532 Human GPC	C 643	18.6	44.3	442	14	ACL54456	ACL54456 Human col
C 571	18.6	44.3	172	8	ABT14728	Abt14728 Human Bcl	C 644	18.6	44.3	448	8	ABX43134	Abx43134 Bovine ES
C 572	18.6	44.3	221	12	ADG99151	Adg99151 Kidney di	C 645	18.6	44.3	469	4	AAK62772	Aak62772 Human imm
C 573	18.6	44.3	354	3	AAC09559	Aac09559 Human sec	C 646	18.6	44.3	484	9	ACH29613	Ach29613 Human adu
C 574	18.6	44.3	376	4	AAK60572	Aak60572 Human imm	C 647	18.6	44.3	499	14	ACL58600	ACL58600 Human col
C 575	18.6	44.3	376	4	AAK78056	Aak78056 Human imm	C 648	18.6	44.3	576	2	AAQ83892	Aaq83892 Hepatitis
C 576	18.6	44.3	395	9	ACH49830	Ach49830 Human leu	C 649	18.6	44.3	576	2	AAQ83892	Aaq83892 Hepatitis
C 577	18.6	44.3	411	4	AAI81107	Aai81107 Human pol	C 650	18.6	44.3	677	10	ADG25172	Adg25172 HCV plasm
C 578	18.6	44.3	424	9	ACH45339	Ach45339 Human foe	C 651	18.6	44.3	721	3	AAF09548	Aaf09548 Fusarium
C 579	18.6	44.3	487	9	ACH39397	Ach39397 Human foe	C 652	18.6	44.3	721	13	ADU53589	Adu53589 Fusarium
C 580	18.6	44.3	500	9	ACH15873	Ach15873 Human adu	C 653	18.6	44.3	721	14	ADZ91592	Adz91592 Fusarium
C 581	18.6	44.3	533	13	ACN56036	Acn56036 Cotton an	C 654	18.6	44.3	751	4	AAI96348	Aai96348 Human neu
C 582	18.6	44.3	580	2	AAQ78105	Aaq78105 Hepatitis	C 655	18.6	44.3	830	2	AAZ16126	Aaz16126 Human gen
C 583	18.6	44.3	590	10	ADD333983	Add333983 Mouse mit	C 656	18.6	44.3	895	12	ADJ41762	Adj41762 Plant cDN
C 584	18.6	44.3	603	6	ABK73507	Abk73507 Bacillus	C 657	18.6	44.3	957	5	AAS90478	Aas90478 DNA encod
C 585	18.6	44.3	653	13	ADS59248	Ads59248 Bacterial	C 658	18.6	44.3	1629	10	ADC29863	Adc29863 Fertiliza
C 586	18.6	44.3	656	13	ADT18358	Adt18358 Plant cDN	C 659	18.6	44.3	1642	4	AAS01169	Aas01169 Fertiliza
C 587	18.6	44.3	725	13	ADR13190	Adr13190 Human can	C 660	18.6	44.3	1742	4	AAH17569	Aah17569 Human CDN
C 588	18.6	44.3	741	13	ADR63989	Adr63989 Cotton cD	C 661	18.6	44.3	1749	4	AAS01158	Aas01158 Fertiliza
C 589	18.6	44.3	822	11	ACN88277	Acn88277 Breast ca	C 662	18.6	44.3	1749	10	ADC29841	Adc29841 Fertiliza
C 590	18.6	44.3	934	14	AEB66366	Aeb66366 Rice geno	C 663	18.6	44.3	1758	4	AAK83400	Aak83400 Human imm
C 591	18.6	44.3	1040	4	AAK69263	Aak69263 Human imm	C 664	18.6	44.3	1794	10	ADC29839	Adc29839 Fertiliza
C 592	18.6	44.3	1194	4	ABLI13841	Abli13841 Drosophil	C 665	18.6	44.3	1813	13	ADX54110	Adx54110 Plant ful
C 593	18.6	44.3	1236	8	ACA54395	Aca54395 Prokaryot	C 666	18.6	44.3	1897	13	ADU07850	Adu07850 DNA seque
C 594	18.6	44.3	1494	4	AAL06935	Aal06935 Human rep	C 667	18.6	44.3	1930	4	AAK83401	Aak83401 Human imm
C 595	18.6	44.3	1496	4	AAL03200	Aal03200 Human rep	C 668	18.6	44.3	2025	12	ADL03940	Adl03940 DNA encod
C 596	18.6	44.3	1496	4	AAL06941	Aal06941 Human rep	C 669	18.6	44.3	2223	14	AED12837	Aed12837 GH74- end
C 597	18.6	44.3	1496	4	AAL06934	Aal06934 Human rep	C 670	18.6	44.3	2233	14	AED12900	Aed12900 A. cellu
C 598	18.6	44.3	1500	3	AAZ29594	Aaz29594 Japanese	C 671	18.6	44.3	2388	4	ABL30221	Ab130221 Drosophil
C 599	18.6	44.3	1642	12	ADK65901	Adk65901 Potato sa	C 672	18.6	44.3	2446	10	AAD63603	Aad63603 Maize sca
C 600	18.6	44.3	1908	13	ADS61606	Ads61606 Bacterial	C 673	18.6	44.3	2869	10	ADD29922	Add29922 Acidother
C 601	18.6	44.3	2046	13	ADS57707	Ads57707 Bacterial	C 674	18.6	44.3	2869	10	ABZ77632	Abz77632 Nucleotid
C 602	18.6	44.3	2084	13	ADT45082	Adt45082 Bacterial	C 675	18.6	44.3	2869	12	ADO52314	Ado52314 Acidother

c 676	18.4	43.8	3197	10	ADD11037	Adi13037 A. gossyp	749	18.2	43.3	492	4	AAK67080	Human imm
c 677	18.4	43.8	3888	3	AAc47167	Arabidops	750	18.2	43.3	502	4	ABA61924	Human foe
c 678	18.4	43.8	4283	12	ADJ38145	Arabidops	751	18.2	43.3	502	4	AAI41846	Probe #10
c 679	18.4	43.8	4635	4	ABL30220	Drosophil	752	18.2	43.3	502	4	AAK36132	Human bon
c 680	18.4	43.8	4950	10	ADB53099	Primary r	753	18.2	43.3	502	4	AAK10233	Human bra
c 681	18.4	43.8	4950	13	ADV41301	Rat cardi	754	18.2	43.3	502	4	ABS35825	Human liv
c 682	18.4	43.8	6060	12	ADJ38148	Arabidops	755	18.2	43.3	502	6	ABS10268	Human gen
c 683	18.4	43.8	6336	6	ABU67080	Thyroid c	756	18.2	43.3	508	14	ABA38476	Tomato th
c 684	18.4	43.8	6336	6	ABN95039	Gene #153	757	18.2	43.3	566	10	ADK58514	Plant DNA
c 685	18.4	43.8	6336	14	ADJ18235	Fibrotic	758	18.2	43.3	647	13	ADR65375	Cotton cd
c 686	18.4	43.8	6373	2	AAV74538	Staphyloc	759	18.2	43.3	647	13	ADT16840	Plant cDN
c 687	18.4	43.8	6400	12	ADJ38276	Plastid d	760	18.2	43.3	704	13	ADX47448	Plant ful
c 688	18.4	43.8	7525	4	AA501192	Fertiliza	761	18.2	43.3	743	4	AAH05826	Human cDN
c 689	18.4	43.8	7525	10	ADC29908	Fertiliza	762	18.2	43.3	769	13	ADH08351	Plant ful
c 690	18.4	43.8	9893	4	AAK85753	Human imm	763	18.2	43.3	776	4	AAI95091	Human neu
c 691	18.4	43.8	10580	4	AAK69427	Human imm	764	18.2	43.3	801	3	AAc46622	Zea mays
c 692	18.4	43.8	13031	11	ADK39442	Maize Zmf	765	18.2	43.3	811	4	AAI94653	Human neu
c 693	18.4	43.8	15270	4	AAK74017	Human imm	766	18.2	43.3	811	8	ACA56923	Human adi
c 694	18.4	43.8	15270	4	AAK85754	Human imm	767	18.2	43.3	811	13	ADO82213	Plant ful
c 695	18.4	43.8	32191	4	AAU03616	Human rep	768	18.2	43.3	831	8	ACA34548	Prokaryot
c 696	18.4	43.8	32191	4	ABA07814	Human ova	769	18.2	43.3	983	13	ADS45289	Bacterial
c 697	18.4	43.8	41554	11	ACN44900	Mouse gen	770	18.2	43.3	1008	6	ABN67751	Streptoco
c 698	18.4	43.8	42979	4	ABL20870	Drosophil	771	18.2	43.3	1010	2	AAK13969	H. pylori
c 699	18.4	43.8	80336	3	AAE22283	BAC conta	772	18.2	43.3	1011	13	ADV85016	Streptoco
c 700	18.4	43.8	90336	3	AAE22289	BAC conta	773	18.2	43.3	1272	4	AAE54079	hFIX gene
c 701	18.4	43.8	96589	9	AAU57700	Mouse pro	774	18.2	43.3	1272	4	AAE54080	hFIX gene
c 702	18.4	43.8	96592	9	ADA02669	Mouse Pr1	775	18.2	43.3	1273	4	AAE54074	hFIX gene
c 703	18.4	43.8	96592	10	ADB72407	Mouse Pr1	776	18.2	43.3	1273	4	AAE54074	hFIX gene
c 704	18.4	43.8	96592	10	ADE95917	Mouse Pr1	777	18.2	43.3	1273	4	AAE54075	hFIX gene
c 705	18.4	43.8	99629	4	AAE28850	Genomic f	778	18.2	43.3	1273	4	AAE54076	hFIX gene
c 706	18.4	43.8	103661	12	ADQ97712	Mouse can	779	18.2	43.3	1273	4	AAE54073	hFIX gene
c 707	18.4	43.8	106344	10	ADJ79961	Full leng	780	18.2	43.3	1275	4	AAE54077	hFIX gene
c 708	18.4	43.8	110000	6	ABA901993	Continuation (4 of	781	18.2	43.3	1276	4	AAE54078	hFIX gene
c 709	18.4	43.8	110000	6	ABQ69245_07	Continuation (8 of	782	18.2	43.3	1400	14	AAE14143	Transgene
c 710	18.4	43.8	110000	6	ABQ67197_06	Continuation (7 of	783	18.2	43.3	1427	5	AAc82742	C. sinens
c 711	18.4	43.8	110000	6	ABQ87681_3	Continuation (4 of	784	18.2	43.3	1427	5	AAc82743	C. sinens
c 712	18.4	43.8	110000	8	ABX3177_3	Continuation (4 of	785	18.2	43.3	1427	6	AAK99774	1427nt DN
c 713	18.4	43.8	110000	10	ADB81391_2	Continuation (3 of	c 786	18.2	43.3	1493	12	ADO25994	Barley ch
c 714	18.4	43.8	110000	11	ACN44932_2	Continuation (3 of	c 787	18.2	43.3	1496	10	ADB53828	Primary r
c 715	18.4	43.8	110000	11	ACN43998_4	Continuation (5 of	c 788	18.2	43.3	1496	13	ADV41755	Rat cardi
c 716	18.4	43.8	110000	12	ADN46845_07	Continuation (8 of	789	18.2	43.3	1707	6	AAI71008	Human Fac
c 717	18.4	43.8	110000	12	ADN47591_13	Continuation (14 o	790	18.2	43.3	1814	13	ADX09014	Plant ful
c 718	18.4	43.8	110000	12	ADN46123_07	Continuation (8 of	791	18.2	43.3	1814	14	ABE70441	Insulin-1
c 719	18.4	43.8	110000	12	ADN47209_13	Continuation (14 o	792	18.2	43.3	1893	14	ADZ62955	Murine C4
c 720	18.4	43.8	110000	12	ADN46464_07	Continuation (8 of	793	18.2	43.3	1903	13	ADM60050	Plant ful
c 721	18.4	43.8	110000	12	ADN47960_13	Continuation (14 o	c 794	18.2	43.3	1910	11	ADM02479	Human cDN
c 722	18.4	43.8	123192	13	ADV34995	AdV34995 Murine cD	c 795	18.2	43.3	1910	14	AEC85409	Human cDN
c 723	18.4	43.8	135462	12	ADQ97061	Mouse can	c 796	18.2	43.3	1960	4	ABL06493	Drosophil
c 724	18.4	43.8	139257	10	ADC89520	Adg89520 Human COR	c 797	18.2	43.3	1971	4	ABA55251	Human foe
c 725	18.4	43.8	189013	8	ACF62741	Cancer ba	798	18.2	43.3	1971	4	ABS28575	Human liv
c 726	18.4	43.8	189013	8	ADB20856	MRP1 base	799	18.2	43.3	2000	8	ADA72509	Rice gene
c 727	18.4	43.8	189013	10	ADB87945	Human UGT	c 800	18.2	43.3	2000	8	ADA72571	Rice gene
c 728	18.4	43.8	189013	10	ADB96928	Human MDR	c 801	18.2	43.3	2151	10	ACF69072	Photorhab
c 729	18.4	43.8	189013	10	ADB92119	Human MDR	c 802	18.2	43.3	2235	4	AAI92780	Human pol
c 730	18.4	43.8	190242	14	ADZ12571	Murine ca	c 803	18.2	43.3	2283	5	AAE192780	DNA encod
c 731	18.2	43.3	65	6	ABN51456	Mouse spl	c 804	18.2	43.3	2304	8	ACA34716	Prokaryot
c 732	18.2	43.3	122	3	AAI10806	Fusarium	c 805	18.2	43.3	2329	5	AAc90666	DNA encod
c 733	18.2	43.3	122	13	ADU54847	Fusarium	806	18.2	43.3	2352	2	AAI14325	H. pylori
c 734	18.2	43.3	122	14	ADU29850	Adg29850 Fusarium	807	18.2	43.3	2458	14	AAE10501	Human rec
c 735	18.2	43.3	136	7	AD568818	Corn seed	808	18.2	43.3	2481	4	AAH14242	Human cDN
c 736	18.2	43.3	178	8	ABT14715	Human Bcl	809	18.2	43.3	2728	13	ADQ38340	Human SNP
c 737	18.2	43.3	276	5	ABV07935	Human pro	810	18.2	43.3	2758	13	ACN43105	Human dia
c 738	18.2	43.3	297	6	ABL71617	Corn tass	811	18.2	43.3	2766	13	ACN43104	Human dia
c 739	18.2	43.3	329	4	AA337670	Novel hum	812	18.2	43.3	2771	13	ADQ38338	Human SNP
c 740	18.2	43.3	342	8	ACA19445	Pseudomon	813	18.2	43.3	2775	3	AAA54034	Human fac
c 741	18.2	43.3	342	8	ACA19445	Prokaryot	814	18.2	43.3	2777	13	ADQ38339	Human SNP
c 742	18.2	43.3	369	5	ABV37958	Human pro	815	18.2	43.3	2781	1	AAH40177	Part of t
c 743	18.2	43.3	375	11	ABD09197	Pseudomon	816	18.2	43.3	2781	1	AAH40177	Part of t
c 744	18.2	43.3	411	4	AAI89848	Human pol	817	18.2	43.3	2792	2	AAK28626	Nucleotid
c 745	18.2	43.3	454	15	Aef78516	Mouse gen	818	18.2	43.3	2802	2	AAU02460	Human Fac
c 746	18.2	43.3	457	3	AAc44528	Zea mays	819	18.2	43.3	2804	8	ABV77053	Nucleotid
c 747	18.2	43.3	457	3	AAc38313	Zea mays	820	18.2	43.3	2831	13	ADQ38341	Human SNP
c 748	18.2	43.3	469	3	AAc38332	Zea mays	821	18.2	43.3	2880	8	ACC00311	Halohydril

C 822	18.2	43.3	2905	10	ABT41705	Abt41705 Toxicity	895	18.2	43.3	128978	6	ABK83459	Abk83459 Human cDN
C 823	18.2	43.3	2994	5	AAH78232	Aah78232 Nucleotid	896	18.2	43.3	128978	8	AAD54587	Aad54587 Human LIM
C 824	18.2	43.3	3168	6	ABN79811	Abn79811 Fungal 2B	897	18.2	43.3	128978	13	AD52994	Ad52994 Drug ther
C 825	18.2	43.3	3187	12	ADO217407	Ado217407 Human sof	898	18.2	43.3	128978	14	ADX07226	Adx07226 Cyclin-de
C 826	18.2	43.3	3306	3	ADA87701	Aaa87701 Human sec	C 899	18.2	43.3	130414	14	AXD07226	Adx07226 Human CA
C 827	18.2	43.3	3306	8	ADA40549	Ada40549 Human sec	C 900	18.2	43.3	134312	11	ACN44512	Acn44512 Mouse gen
C 828	18.2	43.3	3306	8	ACC50852	Acc50852 Human sec	C 901	18.2	43.3	155462	12	ADQ97394	Adq97394 Mouse can
C 829	18.2	43.3	3306	8	ABZ71477	Abz71477 Secreted	C 902	18.2	43.3	158001	12	ADL17884	Adl17884 Human pho
C 830	18.2	43.3	3306	9	ADB91444	Adb91444 Human sec	C 903	18.2	43.3	161652	12	ADF13118	Adf13118 Hypermech
C 831	18.2	43.3	3306	10	ADC73394	Adc73394 Human sec	C 904	18.2	43.3	161652	12	ADI37264	Adi37264 Hypermech
C 832	18.2	43.3	3306	10	ADA56695	Ada56695 Gene enco	C 905	18.2	43.3	170834	10	AAD62833	Aad62833 Human BAC
C 833	18.2	43.3	3308	3	AAA87681	Aaa87681 Human sec	C 906	18.2	43.3	171936	6	ABS56565	Abs56565 Human SUL
C 834	18.2	43.3	3308	8	ADA40228	Ada40228 Human sec	C 907	18.2	43.3	171936	12	ADN16205	Adn16205 Mouse sul
C 835	18.2	43.3	3308	8	ACC50671	Acc50671 Human sec	C 908	18.2	43.3	171936	12	ADL13775	Adl13775 Osteoarth
C 836	18.2	43.3	3308	8	ABZ71365	Abz71365 Secreted	C 909	18.2	43.3	173805	10	ADL13775	Adl13775 Osteoarth
C 837	18.2	43.3	3308	9	ADB91306	Adb91306 Human sec	C 910	18.2	43.3	196063	13	ABD33462	Abd33462 Human can
C 838	18.2	43.3	3308	10	ADC73394	Adc73394 Human sec	C 911	18.2	43.3	222930	6	ABK84349	Abk84349 Human cDN
C 839	18.2	43.3	3308	10	ADA56388	Ada56388 Gene enco	C 912	18.2	43.3	222930	14	ABD17977	Abd17977 Fibrotic
C 840	18.2	43.3	3397	5	ABV25291	Abv25291 Human pro	C 913	18.2	43.3	243072	10	ACF65382	Acf65382 Photorhab
C 841	18.2	43.3	3799	12	ADO22263	Ado22263 Human sof	C 914	18.2	43.3	252907	13	ABD32694	Abd32694 Human can
C 842	18.2	43.3	3832	6	ABN59890	Abn59890 Novel hum	C 915	18.2	43.3	252907	13	ABD33020	Abd33020 Mouse can
C 843	18.2	43.3	4187	4	ABL29892	AbL29892 Drosophill	C 916	18.2	43.3	259500	15	AEF07224	Aef07224 PFTK1 loc
C 844	18.2	43.3	4268	4	ABL06492	AbL06492 Drosophill	C 917	18.2	43.3	288563	14	ASE05135	Ase05135 Cancer-as
C 845	18.2	43.3	4670	12	ADO35651	Ado35651 Novel mou	C 918	18.2	43.3	349980	6	ABQ81844	Abq81844 Bifidobac
C 846	18.2	43.3	7260	6	ABK35504	Abk35504 Human end	C 919	18.2	42.9	142	60	AAV79046	Aav79046 Staphyloc
C 847	18.2	43.3	7260	6	ABK35561	Abk35561 Gene IGF1	C 920	18.2	42.9	146	8	ABT14725	Abt14725 Human Bcl
C 848	18.2	43.3	7260	6	ABT11091	Abt11091 Human bre	C 921	18.2	42.9	377	13	ADR63349	Adr63349 Cotton cD
C 849	18.2	43.3	7260	6	ABK84583	Abk84583 Human cDN	C 922	18.2	42.9	395	4	AAI85217	Aai85217 Human pol
C 850	18.2	43.3	7260	6	ABK64812	Abk64812 Human ben	C 923	18.2	42.9	400	2	AAV78280	Aav78280 Staphyloc
C 851	18.2	43.3	7260	6	ABN97244	Abn97244 Gene #374	C 924	18.2	42.9	423	8	ACF73809	Acf73809 Staphyloc
C 852	18.2	43.3	7260	9	ADA26452	Ada26452 Human ins	C 925	18.2	42.9	453	8	ABX37646	Abx37646 Bovine ES
C 853	18.2	43.3	7260	10	ADG33110	Adg33110 Human DNA	C 926	18.2	42.9	460	12	ADQ04184	Adq04184 Maize tra
C 854	18.2	43.3	7260	11	ADP65015	Adp65015 Human ins	C 927	18.2	42.9	519	14	ADY65571	Ady65571 S. mangon
C 855	18.2	43.3	7260	14	ADV16941	Adv16941 Human ins	C 928	18.2	42.9	527	12	ACH80203	Ach80203 Human gen
C 856	18.2	43.3	7260	14	ADY16236	Ady16236 DNA enco	C 929	18.2	42.9	535	4	AAF25733	Aaf25733 C. purpur
C 857	18.2	43.3	7260	14	ADY16234	Ady16234 DNA enco	C 930	18.2	42.9	613	14	ADZ61968	Adz61968 Murine Mg
C 858	18.2	43.3	7260	15	AEF87543	Aef87543 Human liv	C 931	18.2	42.9	636	2	AAV77975	Aav77975 Staphyloc
C 859	18.2	43.3	7260	15	AEF87543	Aef87543 Human liv	C 932	18.2	42.9	685	14	ADH71697	Adh71697 Coryneb
C 860	18.2	43.3	7476	13	ABD32850	Abd32850 Human can	C 933	18.2	42.9	703	9	ADB82424	Adb82424 Human cDN
C 861	18.2	43.3	7911	14	ABE70428	Aeb70428 Insulin-1	C 934	18.2	42.9	732	6	ABQ89481	Abq89481 Human pro
C 862	18.2	43.3	7935	14	ABE70427	Aeb70427 Insulin-1	C 935	18.2	42.9	736	4	ABQ89481	Abq89481 Human pro
C 863	18.2	43.3	11022	4	AA341739	Aas41739 Genomic s	C 936	18.2	42.9	738	12	ADI42292	Adi42292 Plant tra
C 864	18.2	43.3	11022	4	AAK80617	Aak80617 Spumaretr	C 937	18.2	42.9	768	4	ABL14975	AbL14975 Drosophil
C 865	18.2	43.3	11693	6	ABN85367	Abn85367 Spumaretr	C 938	18.2	42.9	771	11	ADP19671	Adp19671 Leptospir
C 866	18.2	43.3	11757	6	ABN85366	Abn85366 Spumaretr	C 939	18.2	42.9	777	3	ACA47062	Aac47062 Arabidops
C 867	18.2	43.3	14518	6	AB579053	Ab579053 E. coli C	C 940	18.2	42.9	821	4	AAK75860	Aak75860 Human imm
C 868	18.2	43.3	14518	10	ADH80620	Adh80620 Escherich	C 941	18.2	42.9	852	3	AAA13883	Aaa13883 L. interc
C 869	18.2	43.3	15153	12	ADG93315	Adg93315 DEN3 (Sle	C 942	18.2	42.9	866	3	ACA54853	Aac54853 Arabidops
C 870	18.2	43.3	16291	4	AA341741	Aas41741 Genomic s	C 943	18.2	42.9	868	3	ACA38487	Aac38487 Arabidops
C 871	18.2	43.3	16291	4	AAK80619	Aak80619 Human imm	C 944	18.2	42.9	877	3	ACA41927	Aac41927 Arabidops
C 872	18.2	43.3	17929	10	ADC86000	Adc86000 Human GPC	C 945	18.2	42.9	878	14	ABE89417	Aeb89417 Isolated
C 873	18.2	43.3	20776	4	ABL30040	AbL30040 Drosophil	C 946	18.2	42.9	896	13	ADR63326	Adr63326 Cotton cD
C 874	18.2	43.3	24333	12	ADJ12497	Adj12497 DNA fragm	C 947	18.2	42.9	899	11	ACN85147	Acn85147 Breast ca
C 875	18.2	43.3	24333	12	ADJ12606	Adj12606 DNA fragm	C 948	18.2	42.9	919	3	ACN51838	Acn51838 Arabidops
C 876	18.2	43.3	38059	4	AAE54018	Aaf54018 Human fac	C 949	18.2	42.9	940	6	ABQ88077	Abq88077 Human GPC
C 877	18.2	43.3	38059	6	ABN95627	Abn95627 Gene #212	C 950	18.2	42.9	966	5	AAH68212	Aah68212 C glutami
C 878	18.2	43.3	38059	10	ADD71098	AdD71098 Human coa	C 951	18.2	42.9	1041	3	ACA39294	Aac39294 Arabidops
C 879	18.2	43.3	41104	6	AAJ36260	Aad36260 Human G-p	C 952	18.2	42.9	1041	3	ACA39294	Aac39294 Arabidops
C 880	18.2	43.3	54993	11	ACN45086	Acn45086 Human gen	C 953	18.2	42.9	1041	8	ACA34847	Aca34847 Prokaryot
C 881	18.2	43.3	59560	4	AAK80620	Aak80620 Human imm	C 954	18.2	42.9	1042	3	ACA54852	Aac54852 Arabidops
C 882	18.2	43.3	59560	4	AAK71550	Aak71550 Human imm	C 955	18.2	42.9	1080	8	ACA39731	Aca39731 Prokaryot
C 883	18.2	43.3	68193	12	ADQ97400	Adq97400 Mouse can	C 956	18.2	42.9	1097	2	AAK14425	Aax14425 H. pylori
C 884	18.2	43.3	73507	11	ACN44612	Acn44612 Mouse gen	C 957	18.2	42.9	1129	11	ACN85653	Acn85653 Breast ca
C 885	18.2	43.3	95596	13	ADV87741	Adv87741 Streptoco	C 958	18.2	42.9	1137	4	AAF72044	Aaf72044 Coryneb
C 886	18.2	43.3	95596	13	ADV78994	Adv78994 Streptoco	C 959	18.2	42.9	1200	13	ADT42031	Adt42031 Bacterial
C 887	18.2	43.3	104990	13	ABD32848	Abd32848 Human can	C 960	18.2	42.9	1236	3	AAK36186	Aak36186 Arabidops
C 888	18.2	43.3	110000	3	AAF22303_4	Continuation (5 of	C 961	18.2	42.9	1285	13	ADS60929	AdS60929 Bacterial
C 889	18.2	43.3	110000	4	AAI99682_17	Continuation (18 o	C 962	18.2	42.9	1373	13	ADT15055	Adt15055 Plant cDN
C 890	18.2	43.3	110000	4	AAI99683_17	Continuation (18 o	C 963	18.2	42.9	1419	15	AEF00581	Aef00581 Human des
C 891	18.2	43.3	110000	6	AEF71527_03	Continuation (4 of	C 964	18.2	42.9	1533	15	AEF00579	Aef00579 Human des
C 892	18.2	43.3	110000	10	ACF67367_19	Continuation (20 o	C 965	18.2	42.9	1596	15	AEF00577	Aef00577 Human des
C 893	18.2	43.3	110000	13	ADV81204_03	Continuation (4 of	C 966	18.2	42.9	1902	4	ABL05301	AbL05301 Drosophil
C 894	18.2	43.3	121151	15	ABF74657	Aef74657 Human pol	C 967	18.2	42.9	2028	14	AED20066	Aed20066 Human ova


```
XX WPI; 2003-040540/03.
DR
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1937; 244pp; English.
XX
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 497 BP; 126 A; 130 C; 129 G; 112 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 11; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 344 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 385
RESULT 3
ADT96292/c
ID ADT96292 standard; cDNA; 552 BP.
XX
AC ADT96292;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1799.
XX
XX Colon cancer; T cell; tumour protein; C634S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CV, Lodes MJ, Secrlist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
DR
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1811; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC
```

```
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 154 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 113
RESULT 4
ADT42774/c
ID ADT42774 standard; cDNA; 552 BP.
XX
AC ADT42774;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1811.
XX
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
XX 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CV, Lodes MJ, Secrlist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
DR
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1811; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC
```

```
PS Claim 1; SEQ ID NO 1811; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 154 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 113
RESULT 5
AAZ90046
ID AAZ90046 standard; cDNA; 603 BP.
AC AAZ90046;
XX
DT 09-MAY-2000 (first entry)
XX
DE Hydrophobic domain containing protein clone HP10549 coding sequence.
KW Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW anti-inflammatory; infection; bodily characteristic; ss.
XX
OS Homo sapiens.
XX
PN WO200000506-A2.
XX
PD 06-JAN-2000.
XX
PF 18-JUN-1999; 99WO-JP003242.
XX
PR 26-JUN-1998; 98JP-00180008.
XX
PA (SAGA ) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
PI Kato S, Kimura T;
XX
WPI; 2000-160665/14.
DR P-PSDB; AAY78809.
XX
PT Novel human proteins having hydrophobic domains used for research and
PT diagnostic purposes.
XX
PS Claim 3; Page 88; 117pp; English.
XX
CC This sequence represents the hydrophobic domain containing protein, clone
CC HP10549 coding region. The sequence is isolated from a human stomach
CC cancer cell line. The invention relates to human proteins with
CC hydrophobic domains, the DNA and the cDNA encoding them. The
CC polynucleotides and proteins are predicted to have biological activities
CC which make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals. Suggested activities include
CC nutritional activity (nutritional source or supplement); cytokine and
CC cell proliferation/differentiation activity; immune stimulating (e.g. as
CC vaccines) or suppressing activity (e.g. to treat various immune
CC
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC organ transplantation); haematopoiesis regulating activity (e.g. in
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or cardiac cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
SQ Sequence 603 BP; 133 A; 179 C; 137 G; 154 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 3; Length 603;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 550 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 591
RESULT 6
ABQ58527/c
ID ABQ58527 standard; cDNA; 620 BP.
XX
AC ABQ58527;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:2222.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB ) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
XX
WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
```

CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence of
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX

SQ Sequence 620 BP; 152 A; 154 C; 161 G; 142 T; 0 U; 11 Other;

Query Match 100.0%; Score 42; DB 6; Length 620;
Best Local Similarity 100.0%; Pred. No. 1.3e-07; Length 620;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAGAGCTATGAAGTAAAG 42
|||||
Db 154 GCCACGCTTGCCGAGCCCTTACAAGAGCTATGAAGTAAAG 113

RESULT 7

ABQ59698/c
ID ABQ59698 standard; cDNA; 634 BP.

XX ABQ59698;

XX 02-AUG-2002 (first entry)

XX Human colon cancer related nucleotide sequence SEQ ID NO:3393.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;

XX Genetic analysis; diagnostic; antisense therapy; gene; ss.

XX Homo sapiens.

XX WO200229086-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US030732.

XX 02-OCT-2000; 2000US-0237271P.

XX (FARB) BAYER CORP.

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.

PS Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in-tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise

CC antibodies, and to screen for peptide analogues and antagonists
XX

SQ Sequence 634 BP; 153 A; 154 C; 168 G; 146 T; 0 U; 13 Other;

Query Match 100.0%; Score 42; DB 6; Length 634;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAGAGCTATGAAGTAAAG 42
|||||
Db 158 GCCACGCTTGCCGAGCCCTTACAAGAGCTATGAAGTAAAG 117

RESULT 8

ABT22328/c
ID ABT22328 standard; DNA; 642 BP.

XX AC ABT22328;

XX 16-APR-2003 (first entry)

XX Breast cancer marker gene SEQ ID No 701.

XX Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;
KW chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;
KW surrogate marker gene; pharmacodynamic marker gene; transgenic animal;
KW human; ds.

XX Homo sapiens.

XX WO200285298-A2.

XX 31-OCT-2002.

XX 19-APR-2002; 2002WO-US012612.

XX 20-APR-2001; 2001US-0285163P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;

XX WPI; 2003-093053/08.

XX Novel isolated polypeptide encoded by breast cancer marker gene, useful
PT for diagnosing, staging, monitoring, prognosing and treating diseases
PT associated with breast cancer.

PS Disclosure; Page 187-188; 725pp; English.

XX The invention relates to an isolated polypeptide encoded by a breast
CC cancer marker gene comprising any of 141/21-805 nucleotide sequences,
CC given in the specification. The methods of the invention are useful for
CC diagnosing patients having an identified breast mass or symptoms
CC associated with breast cancer, to diagnose breast cancer or its
CC precursors, and for monitoring the efficacy of treatment of a breast
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also
CC useful for evaluating a patient before, after or during therapy, to
CC evaluate the reduction in a tumour burden. The breast cancer marker gene
CC proteins are useful as immunogens for raising antibodies, by immunising a
CC mammal with a breast cancer marker protein. The marker proteins are
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to
CC identify other proteins which bind to or interact with the marker
CC proteins. The breast cancer marker genes are useful as surrogate marker
CC genes for one or more disorders, disease states or conditions leading to
CC disease states, in particular, breast cancers. The breast cancer marker
CC genes are useful as pharmacodynamic marker genes. An antibody which
CC selectively binds to a protein of a breast cancer marker gene is useful
CC for treating cancers, particularly breast cancers. The host cell of the
CC polynucleotide sequence represents one of the breast cancer marker genes
CC of the invention

SQ Sequence 642 BP; 148 A; 160 C; 180 G; 142 T; 0 U; 12 Other;
Query Match 100.0%; Score 42; DB 10; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGGTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 184 GCCACGGTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 143
|||||
RESULT 9
AD030035
ID AD030035 standard; cDNA; 1074 BP.
AC ADO30035;
XX
XX ADO30035;
DT 29-JUL-2004 (first entry)
XX
DE Human GPCR RA13 polynucleotide, SEQ ID NO:1137.
XX
XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cystostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW viricide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO2004040000-A2.
XX
XX
PD 13-MAY-2004.
XX
XX 09-SEP-2003; 2003WO-US028226.
PF
XX 09-SEP-2002; 2002US-0409303P.
PR
XX 09-APR-2003; 2003US-0461329P.
XX
XX (PRIM-) PRIMAL INC.
PA
XX
XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX
XX
DR WPI; 2004-390329/36.
DR P-PSDB; ADO29632.
XX
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 1137; 542pp; English.
PS
XX
XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising

CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification, those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1074 BP; 215 A; 316 C; 259 G; 284 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 42; DB 12; Length 1074;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGGTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 1018 GCCACGGTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1059
|||||
RESULT 10
AAZ61776
ID AAZ61776 standard; cDNA; 1212 BP.
XX
XX AAZ61776;
AC
XX
XX 27-MAR-2000 (first entry)
DT
XX
XX cDNA encoding human skin cell transmembrane protein, SEQ ID NO:249.
DE
XX
XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumour vascularisation; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
XX Homo sapiens.
OS
XX
XX WO9955865-A1.
PN
XX
XX 04-NOV-1999.
PD
XX
XX 29-APR-1999; 99WO-NZ0000051.
PF
XX
XX 29-APR-1998; 98US-00069726.
PR
XX 09-NOV-1998; 98US-00188930.
PR
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA
XX
XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
PI
XX WPI; 2000-072177/06.
DR
XX Novel polynucleotides useful for the treatment of various conditions
PT including wounds and cancer.
PT
XX
XX Claim 1; Page 155; 235pp; English.
PS
XX
XX The invention relates to novel nucleic acid sequences derived from rat

CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC cells. Polypeptides of the invention may be used to treat inflammation,
CC cancer and neurological diseases. The proteins may be used to stimulate
CC the growth and motility of keratinocytes, to inhibit the growth of cancer
CC cells, to modulate angiogenesis and tumour vascularisation, to modulate
CC skin inflammation, to modulate epithelial cell growth and to inhibit
CC binding of HIV-1 to leukocytes. The invention may also be used to treat
CC growth and developmental defects, skin wounds and hair follicle
CC disorders. Sequences AAZ61606-261832 represent cDNA sequences derived
CC from several mice, rat or human skin cell types. Sequences AAZ61606-
CC Z61649, AAZ61725-261765, AAZ61802-261811 and AAZ61826 encode proteins
CC with an N-terminal signal sequence, indicating that the proteins are
CC secreted. Sequences AAZ61650-261668, AAZ61766-261780, AAZ61812-261817 and
CC AAZ61827-261829 encode proteins with one or more putative transmembrane
CC domains

XX
SQ Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 1.5e-07; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 990 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 11
AAC99709
ID AAC99709 standard; cDNA; 1212 BP.

XX AAC99709;

DT 08-MAR-2001 (first entry)

DE Skin cell cDNA, SEQ ID NO: 249.

XX Human; skin cell; cytostatic; antiinflammatory; anti-HIV; neutropic;
KW neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease; ss.

OS Homo sapiens.

XX WO200069884-A2.

PD 23-NOV-2000.

PF 15-MAY-2000; 2000WO-NZ000075.

PR 14-MAY-1999; 99US-00312283.

PA (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;

PI Murison JG;

XX WPI; 2001-007495/01.

DR P-PSDB; AAB55958.

XX New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases.

PS Claim 1; Page 210-211; 352pp; English.

XX The present polynucleotide encodes a polypeptide which is expressed in
CC mammalian skin cells. The polypeptide is useful for stimulating
CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC tumours, modulating skin inflammation, stimulating the growth of
CC epithelial cells, inhibiting the binding of human immunodeficiency virus

CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
CC neurological diseases. The polynucleotide can be used as a marker, in the
CC identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns

XX Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 1.5e-07; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42

|||||
Db 990 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 12

ABL34861

ID ABL34861 standard; cDNA; 1212 BP.

XX ABL34861;

DT 04-APR-2002 (first entry)

DE Human cDNA isolated from skin cells SEQ ID NO: 249.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnery;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;
KW ss.

OS Homo sapiens.

XX WO200190357-A1.

PD 29-NOV-2001.

PF 24-MAY-2001; 2001WO-NZ000099.

PR 24-MAY-2000; 2000US-0206650P.

PR 25-JUL-2000; 2000US-0221232P.

PA (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

PI Kumble KD;

DR WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.

PS Claim 1; Page 175; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC cDNA of the invention

SQ Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 6; Length 1212;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 990 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031
|||||
ACAA04775
ID ACA04775 standard; cDNA; 1228 BP.
AC ACA04775;
XX
DT 28-MAY-2003 (first entry)
XX
DE cDNA encoding human membrane associated protein fragment #223.
XX
KW Human; ss; gene; microarray; membrane-associated protein; neuropathology;
KW immunopathology; pancreatic disease; cancer; diabetes; hyperlipidaemia;
KW pancreatic cholera; Alzheimer's disease; Huntington's disease; sarcoma;
KW fibrocystic disease; leukaemia; adenocarcinoma; AIDS; allergy; anaemia;
KW asthma; gout; dementia.
XX
OS Homo sapiens.
XX
PN US6492505-B1.
XX
PD 10-DEC-2002.
XX
PF 31-JAN-2000; 2000US-00495050.
XX
PR 01-FEB-1999; 99US-0118318P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Reddy R, Guegler KJ, Au-Young J;
XX WPI; 2003-327324/31.
DR
XX
XX Combination for research/diagnostic applications and for monitoring
PT treatment of e.g., cancer, comprises polynucleotides comprising a
PT fragment of gene encoding membrane-associated proteins, receptors or ion
PT channels.
PS Claim 1; Col 215-216; 147pp; English.
XX
CC The invention relates to a combination comprising several polynucleotide
CC sequences comprising a fragment of gene encoding membrane-associated
CC proteins, receptors or ion channels. The combination is useful as a
CC probe, for research and diagnostic applications, for monitoring the
CC expression of several expressed polynucleotides, in the diagnosis and
CC monitoring of treatment of pancreatic disease, cancer, immunopathology or
CC neuropathology, for investigating an individual's predisposition to the
CC above disease, in genetic or gene expression analysis of polynucleotide
CC sequences, to investigate cellular responses to infection or drug
CC treatment, as hybridisable array elements in a microarray, to purify a
CC subpopulation of mRNAs, cDNAs or genomic fragments, in drug discovery and
CC diagnostics, prognostics and treatment regimens, in drug discovery and
CC development, in toxicological and carcinogenicity studies, and in
CC forensics or pharmacogenomics, to monitor the progression of disease, to
CC monitor the efficacy of treatment, to diagnose the conditions of the
CC pancreas e.g. diabetes, pancreatic cholera, hyperlipidaemia or
CC fibrocystic disease, to diagnose a cancer e.g. leukaemia, adenocarcinoma
CC or sarcoma, to diagnose immunopathologies e.g. AIDS, allergies, anaemia,
CC asthma or gout, to diagnose neuropathologies e.g. Alzheimer's disease,
CC dementia or Huntington's disease, to rapidly screen large numbers of
CC candidate drug molecules and as query sequences against GenBank,
CC SwissProt, BLOCKS and PRINTS databases. The combination is employed to
CC fine tune the treatment regimen and thus the expression patterns
CC associated with undesirable side effects are avoided. The present
CC sequence represents a cDNA encoding a fragment of gene encoding human
CC membrane-associated proteins, receptors or ion channels
XX
SQ Sequence 1228 BP; 272 A; 341 C; 263 G; 332 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 8; Length 1228;

Best Local Similarity 100.0%; Pred. No. 1.5e-07; Mismatches 0; Gaps 0;
Matches 42; Conservative 0; Indels 0;
QY 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 541 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 582
|||||
RESULT 14
AEA00112
ID AEA00112 standard; cDNA; 1460 BP.
XX
AC AEA00112;
XX
DT 28-JUL-2005 (first entry)
XX
DE Human TAT143 cDNA SEQ ID NO:64.
XX
KW ss; gene; tumor-associated antigen; cytostatic; breast tumor;
KW endocrine disease; gynecology and obstetrics; neoplasm; colon tumor;
KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
KW ovary tumor; skin tumor; liver tumor.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 128..1201
FT /*tag= a
FT /product= "TAT143"
XX
PN US2005106644-A1.
XX
PD 19-MAY-2005.
XX
PF 08-SEP-2004; 2004US-00936626.
XX
PR 20-JUN-2001; 2001US-0299500P.
PR 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323268P.
PR 19-JUN-2002; 2002US-00177488.
PR 26-MAR-2004; 2004US-0557116P.
PR 04-AUG-2004; 2004US-0598899P.
XX
PA (GETH) GENENTECH INC.
XX
PI Cairns B, Chen R, Frantz G, Hillan KJ, Koeppe H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Slivkowski M;
XX
XX WPI; 2005-384304/39.
DR P-PSDB; AEA00190.
XX
PT Novel isolated antibody capable of binding to tumor-associated antigenic
PT target polypeptide, useful for treating cell proliferative disorder e.g.
PT cancer.
XX
PS Claim 1; SEQ ID NO 64; 337pp; English.
XX
CC The invention relates to a novel isolated antibody binding to a
CC polypeptide having at least 80% sequence identity to a polypeptide having
CC any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
CC target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
CC specification, a polypeptide having any one of (AEA00127-AEA00202),
CC lacking its associated signal peptide, or an extracellular domain of a
CC polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
CC encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
CC An antibody of the invention has cytostatic activity. The antibody is
CC useful for inhibiting growth of a cell expressing TAT188, which involves
CC contacting the cell with the antibody. The cell is a cancer cell chosen
CC from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
CC liver cell. The cancer cell is a mammalian cell, preferably a human cell.
CC The antibody is also useful for detecting the level of TAT188

CC polypeptide expressed in a test cell relative to a control cell, and for
CC detecting the level of TAT188 polypeptide or a polypeptide having at
CC least 80 % sequence identity to the TAT188 polypeptide sequence in a test
CC cell relative to a control cell. The antibody is useful for treating a
CC preventing a cell proliferative disorder associated with increased
CC expression or activity of a polypeptide having at least 80 % identity to
CC a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
CC The method of the invention is useful for inhibiting the growth of a
CC cancer cell. The present sequence encodes a polypeptide of the invention.
XX
SQ Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 14; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
|||||
DB 1145 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1186

RESULT 15

AEA00632
ID AEA00632 standard; cDNA; 1460 BP.

XX
AC AEA00632;

DT 28-JUL-2005 (first entry)

DE Human TAT143 cDNA sequence SeqID64.

XX antibody identification; tumor-associated antigen; cytostatic;
KW RNA interference; gene therapy; cell death; cancer; breast tumor;
KW colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;
KW skin tumor; liver tumor; gene; ss; TAT143.

XX Homo sapiens.

PN US2005107595-A1.

XX 19-MAY-2005.

PF 10-SEP-2004; 2004US-00938061.

XX 20-JUN-2001; 2001US-0299500P.

PR 29-JUN-2001; 2001US-0301880P.

PR 18-SEP-2001; 2001US-0323268P.

PR 19-JUN-2002; 2002US-00177488.

PR 26-MAR-2004; 2004US-0557116P.

PR 04-AUG-2004; 2004US-0598899P.

XX (GETH) GENENTECH INC.

PA Cairns B, Chen R, Frantz G, Hillan KJ, Koeppen H, Phillips HS;

PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;

PI Sakanaka C, Chuntharapai A, Reed CJ;

XX WPI; 2005-371577/38.

DR P-PSDB; AEA00710.

XX Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds

PT to tumor-associated antigenic target polypeptide, useful for diagnosing

PT or treating cancer.

XX Claim 1; SEQ ID NO 64; 96pp; English.

XX This invention relates to a novel isolated antibody, for example anti-E16

CC or anti-TAT112 antibody, that binds to a tumor-associated antigenic

CC target polypeptide (TAT) and that lacks an associated signal peptide

CC sequence. The invention may be useful for the development of compounds

CC with a cytostatic activity acting as antagonists of the TAT188

CC polypeptide or RNA interference whilst the disclosed sequences may be

CC useful for gene therapy. The invention is useful for inducing the death

CC of a cell (such as a cancer cell chosen from breast, colon, rectum,
CC endometrium, kidney, lung, ovary, skin and liver) to which it binds,
CC inhibiting proliferation or promoting cell death of a cell expressing
CC TAT188. In addition, the invention may be useful for detecting the level
CC of TAT188 polypeptide in a test cell relative to a control cell, or
CC treating or preventing a cell proliferative disorder associated with
CC increased expression of TAT188. The novel antibody of the invention is
CC useful for inhibiting the growth of a cancer cell and may be useful for
CC diagnosing or treating cancer. The present sequence is that of the human
CC TAT143 cDNA which encodes a protein against which an antibody of the
CC invention may be targeted.
XX

SQ Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 14; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
|||||
DB 1145 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1186

RESULT 16

AAF58615
ID AAF58615 standard; cDNA; 1619 BP.

XX AAF58615;

DT 24-APR-2001 (first entry)

DE Human RECAP polynucleotide, SEQ ID NO: 43.

XX Human; RECAP; receptors and associated proteins; cerebroprotective;
KW nontropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW antihypertoid; immunosuppressive; nephrotropic; antitumor; thyromimetic;
KW cytostatic; antibacterial; virucide; fungicide; proteaseinhibitor;
KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.

XX Homo sapiens.

PN WO200107612-A2.

XX 01-FEB-2001.

PF 21-JUL-2000; 2000WO-US020035.

XX 21-JUL-1999; 99US-0145232P.

PR 07-OCT-1999; 99US-0158578P.

PR 12-NOV-1999; 99US-0165192P.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;

PI Baughn MR, Lu DM, Hillman JL, Patterson C, Lal P;

XX WPI; 2001-168554/17.

DR P-PSDB; AAB68891.

XX Novel receptors and associated proteins for diagnosis and treatment of

PT neurological disorders, immunological disorders including autoimmune/

PT inflammatory disorders and cell proliferative disorders such as cancer.

XX Example 5; Page 126-127; 128pp; English.

XX The present sequence encodes a human RECAP (receptors and associated

CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful

CC in the diagnosis, treatment and prevention of neurological disorders such

CC as stroke, Alzheimer's disease, Pick's disease, Huntington's disease,

CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral

CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD

CC (Creutzfeldt-Jakob disease), GSS (Gerstmann -Strausler-Scheinker

```
CC syndrome); immunological disorders, including autoimmune/inflammatory
CC disorders such as AIDS, DiGeorge's syndrome, severe combined
CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,
CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis and cancer
XX
SQ Sequence 1619 BP; 331 A; 463 C; 400 G; 425 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 4; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1137 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1178
RESULT 17
AAZ90056
ID AAZ90056 standard; cDNA; 1718 BP.
AC
AC AAZ90056;
XX
XX 09-MAY-2000 (first entry)
XX
XX Hydrophobic domain containing protein clone HP10549 nucleotide sequence.
DE
XX Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
XX cell proliferation; immune stimulant; immune deficiency; tumour; pain;
XX rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
XX myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
XX anti-inflammatory; infection; bodily characteristic; ss.
OS Homo sapiens.
XX
XX WO200000506-A2.
XX
XX 06-JAN-2000.
XX
XX 18-JUN-1999; 99WO-JP003242.
XX
XX 26-JUN-1998; 98JP-00180008.
XX
XX (SAGA ) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
XX
XX WPI: 2000-160665/14.
XX P-PSDB; AAY78809.
XX
XX Novel human proteins having hydrophobic domains used for research and
XX diagnostic purposes.
XX
XX Claim 4; Page 111-113; 117pp; English.
XX
XX This sequence represents the hydrophobic domain containing protein, clone
XX HP10549 nucleotide sequence. The sequence is isolated from a human
XX stomach cancer cell line. The invention relates to human proteins with
XX hydrophobic domains, the DNA and the cDNA encoding them. The
XX polynucleotides and proteins are predicted to have biological activities
XX which make them suitable for treating, preventing or ameliorating medical
XX conditions in humans and animals. Suggested activities include
XX nutritional activity (nutritional source or supplement); cytokine and
XX cell proliferation/differentiation activity; immune stimulating (e.g. as
XX vaccines) or suppressing activity (e.g. to treat various immune
XX deficiencies such as SCIDs or HIV, connective tissue disease, systemic
XX lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
XX inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC organ transplantation); haematopoiesis regulating activity (e.g. in
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or cardiac cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
XX Sequence 1718 BP; 381 A; 495 C; 377 G; 465 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 3; Length 1718;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 561 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 602
RESULT 18
ABT10173
ID ABT10173 standard; cDNA; 2302 BP.
XX
XX AC ABT10173;
XX
XX 04-DEC-2002 (first entry)
XX
XX Human breast cancer associated coding sequence SEQ ID NO: 307.
DE
XX
XX Human; breast specific gene; breast cancer; differential expression;
XX cytostatic; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2002S9271-A2.
XX
XX 01-AUG-2002.
XX
XX 25-JAN-2002; 2002WO-US002176.
XX
XX 25-JAN-2001; 2001US-0263757P.
XX 25-APR-2001; 2001US-0286090P.
XX 23-MAY-2001; 2001US-0292517P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Orr MS, Nation M, Diggans JC, Zeng W;
XX WPI; 2002-674803/72.
XX
XX Diagnosing breast cancer in a patient comprises detecting the level of
XX gene expression in cell or tissue samples, where a differential gene
XX expression is indicative of breast cancer.
XX
XX Claim 1; SEQ ID NO 307; 260pp + Sequence Listing; English.
XX
XX The present invention relates to methods of diagnosing breast cancer in a
XX patient, which comprise detecting the level of expression in a tissue
XX sample of two or more genes selected from those shown in ABT09867-
XX ABT1112, where a differential expression of the genes indicates breast
XX cancer. The methods are useful in diagnosing, treating, detecting the
XX progression, and in monitoring treatment of breast cancer in patients.
XX The methods are also useful as a screening tool for agents that modulate
XX the onset or progression of breast cancer. The breast cancer genes may be
XX
```

CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub.published_pct_sequences

SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 6; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
|||||
DB 1117 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1158

RESULT 19

ACC58386
ID ACC58386 standard; cDNA; 2302 BP.

XX ACC58386;

XX 26-AUG-2003 (first entry)

XX Human GPCR-like retinoic acid-induced gene 1.

XX Human; retinoic acid-induced gene 1; RAIG1; feeding; fasting; GPCR;
XX receptor; G-protein coupled receptor; anorectic; antidiabetic;
XX antidepressant; immunomodulator; transgenic; gene therapy; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH Key 100.1173
FT CDS /tag= a
FT /product= "GPCR-like RAIG1"

XX WO2003016553-A2.

XX 27-FEB-2003.

XX 20-AUG-2002; 2002WO-US026510.

XX 20-AUG-2001; 2001US-0313940P.

XX (GETH) GENENTECH INC.
XX (CURA-) CURAGEN CORP.

XX Lewin DA, Stewart TA;

XX WPI; 2003-278580/27.

XX P-PSDB; ABR42649.

XX New G-protein coupled receptor-like retinoic acid induced gene 1 (GPCR-
XX like RAIG1) polypeptide and gene, useful for diagnosing or treating
XX metabolic disorders, e.g. obesity, anorexia, cachexia or diabetes.

XX Claim 10; Page 16-17; 150pp; English.

XX The present sequence is that of human G-protein coupled receptor-like
XX retinoic acid induced gene 1 (GPCR-like RAIG1). This is the human
XX homologue of murine GPCR-like RAIG1. The murine gene was shown to be
XX differentially regulated during fasting-feeding cycles in mice, with
XX moderate induction early in fasting, down-regulation with extended
XX fasting and 4-fold up-regulation with feeding in recovery from fasting.
XX The differentially expressed gene, its mRNA, and the encoded protein, can
XX each be manipulated to detect and treat metabolic disorders associated
XX with up- or down-regulation of GPCR-like RAIG1 activity, such as obesity,
XX anorexia, cachexia or diabetes

SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 8; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
|||||
DB 1117 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1158

RESULT 20

ADD93240
ID ADD93240 standard; cDNA; 2302 BP.

XX ADD93240;

XX 29-JAN-2004 (first entry)

XX RAIG1 coding sequence.

XX ss; gene; retinoic acid-inducible gene 1; RAIG1; orphan;
XX G-protein coupled receptor; GPCR; chromosome 12; GPCR5B; carcinoma;
XX vaccine; breast cancer; pancreatic cancer; lung cancer; liver cancer;
XX ovarian cancer; colon cancer; osteosarcoma.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Key 100.1173
FT CDS /tag= a
FT /product= "RAIG1"

XX WO2003087832-A2.

XX 23-OCT-2003.

XX 10-APR-2003; 2003WO-GB001587.

XX 11-APR-2002; 2002GB-00008331.

XX 17-SEP-2002; 2002GB-00021538.

XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

XX Terrett JA;

XX WPI; 2003-845382/78.

XX P-PSDB; ADD93239.

XX GENBANK; AF095448.

XX Screening, diagnosing and/or treating carcinoma, including breast,
XX pancreatic, lung, liver, ovarian and colon cancer by detecting the change
XX in expression or activity of an RAIG1 polypeptide or encoding nucleic
XX acid molecule.

XX Claim 1; Fig 2; 43pp; English.

XX This sequence encodes retinoic acid-inducible gene 1 (RAIG1) polypeptide.
XX RAIG1 is an orphan G-protein coupled receptor (GPCR) located on
XX chromosome 12. RAIG1 shows a restricted expression pattern compared to a
XX related receptor, GPCR5B, which is widely expressed in peripheral and
XX central tissues. The RAIG1 cDNA and polypeptide sequence may be used in
XX the method of the invention for screening for and/or diagnosis of
XX carcinoma in a subject, and/or monitoring the effectiveness of carcinoma
XX therapy. The method comprises detecting and/or quantifying in a
XX biological sample obtained from the subject an RAIG1 polypeptide and a
XX nucleic acid molecule. The RAIG1 polypeptide and nucleic acid molecule
XX are useful in the manufacture of a medicament for the treatment of
XX carcinoma, where the composition is a vaccine. An agent which interacts
XX with or causes change in the expression or activity of an RAIG1
XX polypeptide or nucleic acid molecule, is also useful in the manufacture
XX of a medicament for the treatment of carcinoma that is breast cancer,
XX pancreatic cancer, lung cancer, liver cancer, ovarian cancer, colon

CC cancer and/or osteosarcoma. They can also be used in the diagnosis and
CC screening of such carcinomas.
SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.7e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1117 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158
RESULT 21
ADL24773
ID ADL24773 standard; DNA; 2302 BP.
XX
AC ADL24773;
XX
DT 20-MAY-2004 (first entry)
XX
DE Intestinal epithelium/peyer's patch M cell-associated DNA sequence #171.
XX
KW intestinal epithelium cell development; peyer's patch M cell development;
KW inflammatory bowel disease; glutenenteropathy; infectious disease;
KW autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;
KW Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;
KW immune system disorder; hypersensitivity; anaphylaxis;
KW blood group incompatibility; ds; human.
XX
OS Homo sapiens.
XX
PN WO200280852-A2.
XX
PD 17-OCT-2002.
XX
PF 04-APR-2002; 2002WO-US010873.
XX
PR 04-APR-2001; 2001US-0281416P.
XX
PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
PI Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;
XX WPI; 2003-075470/07.
XX
PT Novel isolated or purified polypeptide encoded by genes associated with
PT intestinal epithelium or M cell development, differentiation or function,
PT useful for treating autoimmune diseases and infectious diseases.
XX
PS Claim 1; SEQ ID NO 283; 152pp; English.
XX
CC The invention comprises DNA sequences which are associated with
CC intestinal epithelium and peyer's patch M cells. The DNA sequences of the
CC invention are useful for assessing, modifying, modulating or regulating
CC intestinal epithelium or M cell development. The DNA sequences of the
CC invention are also useful in the treatment of: inflammatory bowel
CC disease, glutenenteropathy, infectious diseases, autoimmune diseases
CC (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's
CC disease, multiple sclerosis, allergy, asthma and diabetic mellitus),
CC diseases or disorders of the immune system, hypersensitivity,
CC anaphylaxis, and blood group incompatibility. The present nucleic acid
CC represents an intestinal epithelium/peyer's patch M cell-associated DNA
CC sequence of the invention. NOTE: The present sequence is not shown in the
CC specification, but has been retrieved from the WIPO website.
XX
SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1117 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158
RESULT 22
ABT31923
ID ABT31923 standard; DNA; 2316 BP.
XX
AC ABT31923;
XX
DT 01-MAY-2003 (first entry)
XX
DE Human breast cancer / ovarian cancer related coding sequence #30.
XX
KW Human; gene; ds; cytostatic; breast cancer; ovarian cancer.
XX
OS Homo sapiens.
XX
PN WO2003000012-A2.
XX
PD 03-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019773.
XX
PR 21-JUN-2001; 2001US-0300159P.
PR 27-JUN-2001; 2001US-0301351P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Veiby OP;
XX
DR WPI; 2003-267848/26.
DR P-PSDB; ABJ37054.
XX
PT Determining the presence of breast cancer in an individual, involves
PT using specific polynucleotide markers.
XX
PS Disclosure; Page 163; 233pp; English.
XX
CC The invention comprises a method for assessing whether a patient is
CC afflicted with breast cancer or ovarian cancer. The method involves the
CC use of specific DNA markers. The method of the invention is useful in the
CC detection and treatment of ovarian and breast cancer. DNA sequences
CC ABT31894 - ABT31949 encode human breast/ovarian cancer-related proteins
XX
SQ Sequence 2316 BP; 502 A; 670 C; 535 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 10; Length 2316;
Best Local Similarity 100.0%; Pred. No. 1.7e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1123 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1164
RESULT 23
AAH14688
ID AAH14688 standard; cDNA; 2446 BP.
XX
AC AAH14688;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12388.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX

```

PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 12388; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2446 BP; 507 A; 723 C; 572 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 4; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312
RESULT 24
ABZ42832
ID ABZ42832 standard; DNA; 2456 BP.
XX
XX ABZ42832;
AC
XX
XX 04-MAR-2003 (first entry)
DT
XX
XX Human G protein-coupled receptor RAIG1 nucleotide SEQ ID NO:453.
DE
XX
XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;

```

```

KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
OS Homo sapiens.
XX
XX WO200261087-A2.
PN
XX
XX 08-AUG-2002.
XX
XX 19-DEC-2001; 2001WO-US050107.
PF
XX
XX 19-DEC-2000; 2000US-0257144P.
PR
XX
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
PA
XX
XX Burmer GC, Roush CL, Brown JP;
PI
XX
XX WPI; 2003-046718/04.
DR
XX
XX P-PSDB; ABP81984.
DR
XX
XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
XX Disclosure; Fig 1; 523pp; English.
XX
XX The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, or
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312
RESULT 25
ACC72695
ID ACC72695 standard; cDNA; 2456 BP.
XX
XX ACC72695;
AC

```

XX 09-JUL-2003 (first entry)
XX Human cancer related protein encoding cDNA SEQ ID NO:34.
DE Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometriosis; gene; ss.
XX Homo sapiens.
XX WO2003025138-A2.
XX 27-MAR-2003.
XX 17-SEP-2002; 2002WO-US029560.
XX 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-0323887P.
PR 13-NOV-2001; 2001US-0350666P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI Zlotnik A;
XX WPI: 2003-354600/33.
DR P-PSDB; ABR58568.
XX New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX Claim 8; Page 643-644; 767pp; English.
XX The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX SQ Sequence 2456 BP; 522 A; 720 C; 571 G; 643 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 10; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
RESULT 26
ID ADN39301
ID ADN39301 standard; cDNA; 2456 BP.
XX

AC ADN39301;
XX 17-JUN-2004 (first entry)
XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:619.
DE Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
XX WO2003042661-A2.
XX 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 09-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368099P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI: 2003-468649/44.
DR P-PSDB; ADN39302.
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX Claim 8; SEQ ID NO 619; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38693-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a nucleic acid sequence of the invention.


```
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 11; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
DB 1271 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1312

RESULT 27
ADI28525
ID ADI28525 standard; cDNA; 2456 BP.
XX
AC ADI28525;
XX
DT 22-APR-2004 (first entry)
XX
DE Human GPCR retinoic acid induced 3 (RAI-3) cDNA.
XX
KW Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
KW chromosome 12p13-p12.3; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(112,r)
FT /*tag= a
FT /*label= RAI-3-s1
FT /*note= "located in 5' untranslated region"
FT /*standard_name= "Single nucleotide polymorphism"
FT CDS 254..1327
FT /*tag= b
FT /*product= "Human RAI3"
FT variation replace(364,y)
FT /*tag= c
FT /*label= RAI-3-s2
FT /*note= "exon 1, silent (Ala/Ala)"
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(511,y)
FT /*tag= d
FT /*label= RAI-3-s3
FT /*note= "exon 2, silent (Ile/Ile)"
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(523,y)
FT /*tag= e
FT /*label= RAI-3-s4
FT /*note= "exon 2, silent (Asp/Asp)"
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(605,r)
FT /*tag= f
FT /*label= RAI-3-s6
FT /*note= "exon 2, missense (Ser/Gly)"
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(797,r)
FT /*tag= g
FT /*label= RAI-3-s5
FT /*note= "exon 2, missense (Thr/Ala)"
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(1111,r)
FT /*tag= h
FT /*label= RAI-3-s8
FT /*note= "silent (Pro/Pro)"
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(1173,r)
FT /*tag= i
FT /*label= RAI-3-s9
FT /*note= "missense (Gln/Arg)"
FT /*standard_name= "Single nucleotide polymorphism"
XX
```

PN WO2004001060-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019255.
XX
PR 20-JUN-2002; 2002US-0390850P.
PR 25-AUG-2002; 2002US-0407006P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Whitney GS, Opitack G, Garulacan L, Ramanathan CS, Mckinnon M;
PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX
DR WPI; 2004-090973/09.
DR P-PSDB; ADI28526.
XX
XX
PT New nucleic acid molecule encoding a human G-protein coupled receptor,
PT RAI3, useful for preventing, treating or ameliorating chronic obstructive
PT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
PT of COPD.
XX
PS Claim 21; SEQ ID NO 18; 301pp; English.
XX
CC The present sequence is that of cDNA encoding a human G-protein coupled
CC receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
CC Proteomics methods were used to isolate cigarette smoke-inducible
CC tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
CC identified as being tyrosine phosphorylated and/or as being
CC associated/complexed with tyrosine phosphorylated proteins only in those
CC cells that had been exposed to cigarette smoke. Since RAI-3 is primarily
CC expressed in lung tissue, and since cigarette smoke is a major causative
CC factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
CC novel cellular target for identifying modulators, e.g. agonists or
CC antagonists, useful for the treatment and/or prevention of COPD and
CC related disorders such as emphysema and chronic bronchitis. RAI-3
CC modulators, e.g. agonists and antagonists, especially antisense
CC compounds, can be used to treat COPD and/or its associated or diseases
CC associated with regulation of NF-kB and/or its associated or interacting
CC signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
CC the RAI-3 gene are useful for determining COPD association in
CC individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
CC preventing, treating or ameliorating disorders related to aberrant GPCR
CC signaling or cell cycle regulation, pulmonary disorders, inflammatory
CC lung disorders, COPD, the underlying symptoms of COPD, COPD-related
CC disorders or conditions, autoimmune disorders, disorders related to
CC hyperimmune activity, inflammatory conditions, disorders related to
CC aberrant acute phase responses, hypercongenital conditions, birth
CC defects, necrotic lesions, wounds, organ transplant rejection, renal
CC diseases, ischaemia-reperfusion injury, heart disorders, disorders
CC related to aberrant signal transduction, proliferation disorders,
CC cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
CC ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
CC nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
CC rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
CC pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
CC regulation (all claimed).
XX
SQ Sequence 2456 BP; 516 A; 719 C; 570 G; 643 T; 0 U; 8 Other;

Query Match 100.0%; Score 42; DB 12; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
DB 1271 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1312

RESULT 28

```
ADI28459
ID ADI28459 standard; cDNA; 2456 BP.
XX
AC ADI28459;
XX
DT 22-APR-2004 (first entry)
XX
DE Human GPCR retinoic acid induced 3 (RAI-3) cDNA.
XX
KW Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
KW chromosome 12p13-p12.3; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(112,a)
FT /tag= a
FT /label= RAI-3-s1
FT /note= "located in 5' untranslated region"
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 254..1327
FT /tag= b
FT /product= "Human RAI3"
FT variation replace(364,t)
FT /tag= c
FT /label= RAI-3-s2
FT /note= "exon 1, silent (Ala/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(511,t)
FT /tag= d
FT /label= RAI-3-s3
FT /note= "exon 2, silent (Ile/Ile)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(523,t)
FT /tag= e
FT /label= RAI-3-s4
FT /note= "exon 2, silent (Asp/Asp)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(605,g)
FT /tag= f
FT /label= RAI-3-s6
FT /note= "exon 2, missense (Ser/Gly)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(797,g)
FT /tag= g
FT /label= RAI-3-s5
FT /note= "exon 2, missense (Thr/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(1111,g)
FT /tag= h
FT /label= RAI-3-s8
FT /note= "silent (Pro/Pro)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(1173,g)
FT /tag= i
FT /label= RAI-3-s9
FT /note= "missense (Gln/Arg)"
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO2004001060-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019255.
XX
PR 20-JUN-2002; 2002US-0390850P.
PR 29-AUG-2002; 2002US-0407006P.
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX Whitney GS, Opiteck G, Garulacan L, Ramanathan CS, McKinnon M;
PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
WPI; 2004-090973/09.
P-PSDB; ADI28460.
GENEANK; NM_003979.
New nucleic acid molecule encoding a human G-protein coupled receptor,
RAI3, useful for preventing, treating or ameliorating chronic obstructive
pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
of COPD.
Claim 1; SEQ ID NO 2; 301pp; English.
The present sequence is that of cDNA encoding a human G-protein coupled
receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
Proteomics methods were used to isolate cigarette smoke-inducible
tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
identified as being tyrosine phosphorylated and/or as being
associated/complexed with tyrosine phosphorylated proteins only in those
cells that had been exposed to cigarette smoke. Since RAI-3 is primarily
expressed in lung tissue, and since cigarette smoke is a major causative
factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
novel cellular target for identifying modulators, e.g. agonists or
antagonists, useful for the treatment and/or prevention of COPD and
related disorders such as emphysema and chronic bronchitis. RAI-3
modulators, e.g. agonists and antagonists, especially antisense
compounds, can be used to treat COPD and other disorders and diseases
associated with regulation of NF-kB and/or its associated or interacting
signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
the RAI-3 gene are useful for determining COPD association in
individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
preventing, treating or ameliorating disorders related to aberrant GPCR
signaling or cell cycle regulation, pulmonary disorders, inflammatory
lung disorders, COPD, the underlying symptoms of COPD, COPD-related
disorders or conditions, autoimmune disorders, disorders related to
hyperimmune activity, inflammatory conditions, disorders related to
aberrant acute phase responses, hypercongenital conditions, birth
defects, necrotic lesions, wounds, organ transplant rejection, renal
diseases, ischaemia-reperfusion injury, heart disorders, disorders
related to aberrant signal transduction, proliferation disorders,
cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
CC pneumonia small cell undifferentiated carcinoma, carcinoma of the breast,
CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
CC regulation (all claimed).
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 12; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
RESULT 29
ADR48221
ID ADR48221 standard; cDNA; 2456 BP.
XX
AC ADR48221;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human retinoic acid induced 3 (RAI3) encoding cDNA SEQ:9.
XX
KW pancreatic cancer-associated transcript; pancreatic cancer; human;
KW cytostatic; gene therapy; protein therapy; retinoic acid induced 3; RAI3;
```



```
XX DT 18-NOV-2004 (first entry)
XX DE Human breast tumour associated gene clone-02 SEQ ID NO:2.
XX PD
XX PF cancer; tumour; breast cancer; lung cancer; colon cancer; kidney cancer;
KW antibody; cytostatic; gene therapy; human; breast tumour tissue;
KW breast tissue; gene; ds.
XX OS Homo sapiens.
XX PN WO2004074506-A2.
XX PD
XX PF 02-SEP-2004.
XX PF 30-JAN-2004; 2004WO-US003030.
XX PF 13-FEB-2003; 2003US-0447900P.
XX PA (MERC-) MERGEN LTD.
XX PI Hu Q, Peng A, Liu B, Love JR, Hao X, Ren M, Sheng Z;
XX WI WPI; 2004-635589/61.
XX DR
XX PT Diagnosing and treating a cancer, e.g. breast, lung, colon, or kidney
PT cancer, comprises detecting and blocking the over expression of a gene of
PT a protein found in breast, lung, colon, or kidney tissue.
XX PS Claim 2; SEQ ID NO 2; 220pp; English.
XX CC The present invention describes a method for diagnosing and treating a
CC cancer, e.g. breast, lung, colon, or kidney cancer. The method comprises
CC detecting and blocking the over expression of a gene of a protein found
CC in breast, lung, colon, or kidney tissue. Also described is an antibody
CC or binding portion of an antibody that specifically binds a protein found
CC in breast tissue, lung tissue, colon tissue, or kidney tissue. The
CC antibody has cytostatic activity, and can be used in gene therapy. The
CC methods, antibodies, polynucleotides and polypeptides from the present
CC invention are useful for detecting, diagnosing, preventing and treating
CC cancer, e.g. breast, lung, colon, or kidney cancer. The present sequence
CC represents a nucleotide sequence given in the present invention, which is
CC over expressed in human breast tumour tissue. N.B. All 385 sequences
CC referenced in this patent are detailed in the US provisional application
CC SN 60/447,900, filed 02/13/2003. In this application only one sequence,
CC representing the longest sequence of each of the 65 clones is listed in
CC Appendix A (and given as SEQ ID NO:1 to 65 in the Sequence Listing).
XX QY 1 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1271 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 32
ADU06126
ID ADU06126 standard; DNA; 2456 BP.
XX AC
XX AC ADU06126;
XX DT 27-JAN-2005 (first entry)
XX DE Novel bronchial cancer-associated human gene SeqID350.
XX KW bronchial cancer; cytostatic; tumour-associated protein;
KW cancer detection; metastasis; tumour; gene; ds; human.
XX OS Homo sapiens.
```

```
XX DE10316701-A1.
XX PD
XX PF 04-NOV-2004.
XX PF 09-APR-2003; 2003DE-01016701.
XX PF 09-APR-2003; 2003DE-01016701.
XX PA (HINZ/) HINZMANN B.
XX PA (HERM/) HERMANN K.
XX PA (CAST/) HEIDEN CASTANOS-VELEZ E.
XX PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
XX PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;
XX WI WPI; 2004-786403/78.
XX DR P-PSDB; ADU06613.
XX PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT cancer and in screening for therapeutic and diagnostic agents.
XX PS Claim 1; SEQ ID NO 350; 1381pp; German.
XX CC This invention relates to a novel isolated nucleic acid associated with
CC bronchial cancer comprising 489 defined sequences given in the
CC specification. The invention may be useful for the production of
CC compounds with a cytostatic activity through the inhibition of expression
CC or activity of tumour-associated proteins. The novel DNA sequences and
CC the proteins/peptides encoded by them are used for detecting bronchial
CC cancer or determining the risk of developing it and to screen for
CC specific binding partners of the DNA or protein sequences, where the
CC binding partners are potentially useful as agents for treating or
CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a novel
CC bronchial cancer-associated human gene sequence of the invention.
XX QY Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
XX DB 1271 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

Query Match 100.0%; Score 42; DB 13; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1271 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 33
AEA00088
ID AEA00088 standard; cDNA; 2456 BP.
XX AC
XX AC AEA00088;
XX DT 28-JUL-2005 (first entry)
XX DE Human TAT115 cDNA SEQ ID NO:40.
XX KW ss; gene; tumor-associated antigen; cytostatic; breast tumor;
KW endocrine disease; gynecology and obstetrics; neoplasm; colon tumor;
KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
KW ovary tumor; skin tumor; liver tumor.
XX OS Homo sapiens.
XX FH Key
XX FT 254..1327
XX FT CDS /*tag= a /product= "TAT115"
```

```
XX US2005106644-A1.
PN
XX 19-MAY-2005.
PD
XX
XX 08-SEP-2004; 2004US-00936626.
XX
XX 20-JUN-2001; 2001US-0299500P.
XX
XX 29-JUN-2001; 2001US-0301880P.
XX
XX 18-SEP-2001; 2001US-0323268P.
XX
XX 19-JUN-2002; 2002US-00177488.
XX
XX 26-MAR-2004; 2004US-0557116P.
XX
XX 04-AUG-2004; 2004US-0598899P.
XX
XX (GETH ) GENENTECH INC.
PA
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sliwowski M;
XX
XX WPI; 2005-384304/39.
DR P-PSDB; AEA00166.
DR
XX
XX Novel isolated antibody capable of binding to tumor-associated antigenic
PT target polypeptide, useful for treating cell proliferative disorder e.g.
PT cancer.
PT
XX
XX Claim 1; SEQ ID NO 40; 337pp; English.
XX
XX The invention relates to a novel isolated antibody binding to a
CC polypeptide having at least 80% sequence identity to a polypeptide having
CC any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
CC target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
CC specification, a polypeptide having any one of (AEA00127-AEA00202),
CC lacking its associated signal peptide, or an extracellular domain of a
CC polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
CC encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
CC An antibody of the invention has cytostatic activity. The antibody is
CC useful for inhibiting growth of a cell expressing TAT188, which involves
CC contacting the cell with the antibody. The cell is a cancer cell chosen
CC from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
CC liver cell. The cancer cell is a mammalian cell, preferably a human cell.
CC The antibody is also useful for detecting the level of TAT188
CC polypeptide expressed in a test cell relative to a control cell, and for
CC detecting the level of TAT188 polypeptide or a polypeptide having at
CC least 80 % sequence identity to the TAT188 polypeptide sequence in a test
CC cell relative to a control cell. The antibody is useful for treating a
CC preventing a cell proliferative disorder associated with increased
CC expression or activity of a polypeptide having at least 80 % identity to
CC a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
CC The method of the invention is useful for inhibiting the growth of a
CC cancer cell. The present sequence encodes a polypeptide of the invention.
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 42; DB 14; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312
RESULT 34
AEA00608
ID AEA00608 standard; cDNA; 2456 BP.
XX
AC AEA00608;
XX
XX 28-JUL-2005 (first entry)
DT
XX Human TAT115 cDNA sequence SeqID40.
DE
XX
XX antibody identification; tumor-associated antigen; cytostatic;
KW RNA interference; gene therapy; cell death; cancer; breast tumor;
KW colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;
KW skin tumor; liver tumor; gene; ss; TAT115.
XX
OS Homo sapiens.
XX
XX US2005107595-A1.
PN
XX
XX 19-MAY-2005.
PD
XX
XX 10-SEP-2004; 2004US-00938061.
XX
XX 20-JUN-2001; 2001US-0299500P.
XX
XX 29-JUN-2001; 2001US-0301880P.
XX
XX 18-SEP-2001; 2001US-0323268P.
XX
XX 19-JUN-2002; 2002US-00177488.
XX
XX 26-MAR-2004; 2004US-0557116P.
XX
XX 04-AUG-2004; 2004US-0598899P.
XX
XX (GETH ) GENENTECH INC.
PA
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sakanaka C, Chuntharapai A, Reed CJ;
XX
XX WPI; 2005-371577/38.
DR P-PSDB; AEA00686.
DR
XX
XX Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds
PT to tumor-associated antigenic target polypeptide, useful for diagnosing
PT or treating cancer.
PT
XX
XX Claim 1; SEQ ID NO 40; 96pp; English.
XX
XX This invention relates to a novel isolated antibody, for example anti-E16
CC or anti-TAT112 antibody, that binds to a tumor-associated antigenic
CC target polypeptide (TAT) and that lacks an associated signal peptide
CC sequence. The invention may be useful for the development of compounds
CC with a cytostatic activity acting as antagonists of the TAT188
CC polypeptide or RNA interference whilst the disclosed sequences may be
CC useful for gene therapy. The invention is useful for inducing the death
CC of a cell (such as a cancer cell chosen from breast, colon, rectum,
CC endometrium, kidney, lung, ovary, skin and liver) to which it binds,
CC inhibiting proliferation or promoting cell death of a cell expressing
CC TAT188. In addition, the invention may be useful for detecting the level
CC of TAT188 polypeptide in a test cell relative to a control cell, or
CC treating or preventing a cell proliferative disorder associated with
CC increased expression of TAT188. The novel antibody of the invention is
CC useful for inhibiting the growth of a cancer cell and may be useful for
CC diagnosing or treating cancer. The present sequence is that of the human
CC TAT115 cDNA which encodes a protein against which an antibody of the
CC invention may be targeted.
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 42; DB 14; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312
RESULT 35
AEA047480
ID AEA047480 standard; DNA; 2456 BP.
XX
XX AEA047480;
XX
XX 15-DEC-2005 (first entry)
DT
```

XX DE Retinoic acid-inducible G-protein coupled receptor 3 gene, SEQ ID 2.
XX KW RNA Interference; gene silencing;
KW retinoic acid-inducible G-protein coupled receptor 3; breast tumor;
KW cytostatic; cell growth; gene; ds.
XX OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 254..1327
FT /tag= a
FT /product= "Retinoic acid-inducible G-protein coupled
FT receptor 3"
FT misc_feature 338..403
FT /tag= b
FT /label= Transmembrane domain
FT misc_feature 452..514
FT /tag= c
FT /label= Transmembrane domain
FT misc_feature 497..1057
FT /tag= d
FT /label= 7 transmembrane receptor
FT misc_feature 548..601
FT /tag= e
FT /label= Transmembrane domain
FT misc_feature 656..718
FT /tag= f
FT /label= Transmembrane domain
FT misc_feature 776..832
FT /tag= g
FT /label= Transmembrane domain
FT misc_feature 887..946
FT /tag= h
FT /label= Transmembrane domain
FT polyA_signal 2403..2408
FT /tag= i
FT polyA_site 2442..2456
FT /tag= j
XX JP2005287456-A.
XX
XX 20-OCT-2005.
XX
XX 02-APR-2004; 2004JP-00110232.
XX
XX 02-APR-2004; 2004JP-00110232.
XX (NIPK) NIPPON KAYAKU KK.
XX (UYNI-) UNIV NIPPON IKA.
XX Nagahata T, Emi M;
XX WPI: 2005-737982/76.
XX P-PSDB; AED47479.
XX
XX Substance that suppresses expression of retinoic acid-inducible G-protein
XX coupled receptor-3 RA13 gene or nucleic acid, useful for suppressing cell
XX growth, and as pharmaceutical or therapeutic agent of breast cancer.
XX
XX Claim 1; SEQ ID NO 2; 22pp; Japanese.
XX
XX The invention relates to a novel substance which suppresses the
XX expression of a retinoic acid-inducible G-protein coupled receptor (RA1)3
XX gene. The RA13 gene suppressing agent is useful for suppressing cell
XX growth, and as a pharmaceutical or therapeutic agent of breast cancer,
XX and is also useful in identifying a cell growth inhibitory substance. The
XX RA13 gene suppressing agent enables the identification of a cell growth
XX inhibitory substance. This polynucleotide sequence represents the
XX retinoic acid-inducible G-protein coupled receptor 3 gene of the
XX invention.
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 14; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1271 GCCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
RESULT 36
ABQ54954
ID ABQ54954 standard; cDNA; 2593 BP.
XX AC ABQ54954;
XX DT 22-AUG-2002 (first entry)
XX DE Human ovarian antigen HSKJC61 cDNA, SEQ ID NO:834.
XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.
XX Homo sapiens.
OS
XX WO200200677-A1.
XX PN 03-JAN-2002.
PD
XX 07-JUN-2001; 2001WO-US018569.
PF
XX 07-JUN-2000; 2000US-0209467P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Birse CE, Rosen CA;
XX WPI: 2002-147878/19.
DR P-PSDB; ABP41877.
DR
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
XX
XX Claim 1; SEQ ID NO 834; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders

CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2593 BP; 591 A; 730 C; 605 G; 667 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 6; Length 2593;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1383 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1424

RESULT 37
AAH72766
ID AAH72766 standard; cDNA; 3371 BP.

XX AC AAH72766;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 4040.

XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000WO-US033312.

XX PR 08-DEC-1999; 99US-0169681P.

XX PR 21-DEC-1999; 99US-0171350P.

XX PR 14-MAR-2000; 2000US-0189315P.

XX PR 12-MAY-2000; 2000US-0203791P.

XX PR 09-JUN-2000; 2000US-0210600P.

XX PR 21-JUL-2000; 2000US-0220114P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

XX PS WPI; 2001-375006/39.

XX PT New isolated nucleic acid for diagnosing and treating cervical cancer and
XX for assessing and detecting compounds for treating the cancer.

XX PS Claim 1; Page 840; 1051pp; English.

XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful; to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy

XX SQ Sequence 3371 BP; 741 A; 902 C; 779 G; 946 T; 0 U; 3 Other;

Query Match 100.0%; Score 42; DB 4; Length 3371;

Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1184 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1225

RESULT 38

ACN89274
ID ACN89274 standard; DNA; 4239 BP.

XX AC ACN89274;

XX DT 02-DEC-2004 (first entry)

XX DE Breast cancer related marker, seq id 10424.

XX KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX OS Homo sapiens.

XX PN US2003099974-A1.

XX PD 29-MAY-2003.

XX PF 18-JUL-2002; 2002US-00198846.

XX PR 18-JUL-2001; 2001US-0306220P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX DR WPI; 2003-787014/74.

XX PT Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.

XX PS Disclosure; SEQ ID NO 10424; 36pp; English.

XX CC The invention relates to an isolated polypeptide (I) associated with
XX breast cancer which is encoded by a nucleic acid molecule comprising a
XX nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX the polypeptide of the invention. The activity of the polypeptide of the
XX invention may be described as cytostatic. The antibody is useful for
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX invention are useful in the detection of breast tumours. (I) is useful as
XX a marker for breast cancer and in breast cancer therapy. Sequences given
XX in records ACN78851-ACN92934 represent nucleic acid markers associated
XX with breast cancer. Note: The sequence listing does not form part of the
XX specification but may be obtained in electronic format from the USPTO web
XX site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX SQ Sequence 4239 BP; 859 A; 1161 C; 1075 G; 1141 T; 0 U; 3 Other;

Query Match 100.0%; Score 42; DB 11; Length 4239;

Best Local Similarity 100.0%; Pred. No. 2e-07;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 1679 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1720

RESULT 39

ADX97494
ID ADX97494 standard; DNA; 6730 BP.

XX AC ADX97494;

XX DT 21-APR-2005 (first entry)

```
XX DE Pancreatic cancer associated human protein encoding DNA, SEQ ID 42.
XX KW pancreas tumor; cytostatic; gene; ds.
XX OS Homo sapiens.
XX PN EP1471075-A2.
XX PD 27-OCT-2004.
XX PF 31-MAR-2004; 2004EP-00090124.
XX PR 31-MAR-2003; 2003DE-01015834.
XX PA (HINZ/) HINZMANN B.
XX PA (ROSE/) ROSENTHAL A.
XX PA (PILA/) PILARSKY C.
XX PA (DAHL/) DAHL E.
XX PA (SPEC/) SPECHT T.
XX PA (LICH/) LICHTNER R.
XX PI Rosenthal A, Pilarsky C, Dahl E, Specht T, Bruemendorf T;
XX PI Lichtner R, Staub E, Roepcke S, Li X;
XX DR WPI; 2004-768082/76.
XX DR P-PSDB; ADX97565.
XX PT New nucleic acid differentially expressed in pancreatic tumor tissue, for
XX PT use as diagnostic agents and in screening for therapeutic agents.
XX PS Claim 1; SEQ ID NO 42; 28pp; German.
XX CC The invention relates to a novel human nucleic acid sequence of the
XX CC pancreas and its encoded protein. The invention further comprises:
XX CC proteins and peptides, preferably isolated, that contain a sequence
XX CC encoded by the novel nucleic acid; and methods for diagnosis and
XX CC treatment of pancreatic cancer, using a substance that inhibits or binds
XX CC to the protein or its DNA, including: an antisense oligonucleotide, short
XX CC interfering RNA or ribozyme directed against the pancreatic protein, an
XX CC organic molecule, particularly having a molecular weight below 5000,
XX CC especially 300, that binds to the pancreatic DNA, an aptamer or
XX CC (monoclonal) antibody, preferably human or humanized, that binds to the
XX CC pancreatic DNA, or an anti-idiotypic antibody raised against the
XX CC monoclonal antibody, any of which may be derivatized with a reporter
XX CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
XX CC human pancreatic proteins and their encoding DNA have cytostatic
XX CC activity. The novel sequences are useful for inhibiting transcription
XX CC and/or expression of genes and proteins associated with pancreatic
XX CC cancer. This polynucleotide sequence represents the DNA encoding one of
XX CC the novel human pancreatic proteins of the invention. Note: This sequence
XX CC is not shown in the specification, it has been electronically downloaded
XX CC from a DVD-rom provided with this specification by the European Patent
XX CC Office.
XX SQ Sequence 6730 BP; 1459 A; 1763 C; 1555 G; 1943 T; 0 U; 0 Other;
Query Match 100.0%; Score 42; DB 13; Length 6730;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCGTTGGCGGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1271 GCCCAGCGTTGGCGGCCCTTACAAAGACTATGAAGTAAAG 1312
RESULT 40
ADF70573
ID ADF70573 standard; DNA; 1788 BP.
XX AC ADF70573;
XX DT 12-FEB-2004 (first entry)
```

```
XX DE Orphan receptor ligand-related human protein gene SeqID196.
XX KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
XX KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
XX KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.
XX OS Homo sapiens.
XX PN WO2003071272-A1.
XX PD 28-AUG-2003.
XX PF 21-FEB-2003; 2003WO-JP001901.
XX PR 22-FEB-2002; 2002JP-00045728.
XX PR 23-JUL-2002; 2002JP-00213949.
XX PR 11-OCT-2002; 2002JP-00298237.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
XX DR WPI; 2003-697654/66.
XX DR P-PSDB; ADF70471.
XX PT Transformation of cells with a fusion protein of an orphan receptor
XX PT protein with a fluorescent protein useful for identification of ligands
XX PT to the orphan receptor.
XX PS Example 4; SEQ ID NO 196; 594pp; Japanese.
XX CC This invention relates to a novel method of identifying ligands to an
XX CC orphan receptor protein which comprises transforming cells with DNA
XX CC encoding a fusion protein of the orphan receptor with a fluorescent
XX CC protein, so that the fusion protein is expressed in the cells (or cell
XX CC membranes isolated from them) and contacting the cells with the potential
XX CC ligand to be tested. A suitable fluorescent protein for incorporation in
XX CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
XX CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
XX CC identification of ligands binding to an orphan receptor protein.
XX SQ Sequence 1788 BP; 444 A; 473 C; 412 G; 459 T; 0 U; 0 Other;
Query Match 96.2%; Score 40.4; DB 10; Length 1788;
Best Local Similarity 97.6%; Pred. No. 7.8e-07;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCCAGCGTTGGCGGCCCTTACAAAGACTATGAAGTAAAG 42
DB 1018 GCCCAGCGTTGGCGGCCCTTACAAAGACTATGAAGTAAAG 1059
RESULT 41
AAH08354
ID AAH08354 standard; cDNA; 680 BP.
XX AC AAH08354;
XX DT 26-JUN-2001 (first entry)
DE Human cDNA clone (5'-primer) SEQ ID NO:5189.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
```

```
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 08-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 1; SEQ ID NO 5189; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC polynucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 680 BP; 143 A; 194 C; 151 G; 184 T; 0 U; 8 Other;
SQ
Query Match 80.0%; Score 33.6; DB 4; Length 680;
Best Local Similarity 87.8%; Pred. No. 0.0005;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCGCGAGCCCTTACAAGACTATGAGTAA 41
DB 633 GCCACGCTTGCGCGAGCCCTTACAAGACTATGAGTAA 673
RESULT 42
ADI28530
ID ADI28530 standard; DNA; 42 BP.
XX
XX ADI28530;
XX
XX 22-APR-2004 (first entry)
XX
XX Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA.
XX
XX Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; ss.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX WO2004001060-A2.
XX
XX
```

```
PD 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019255.
XX
XX 20-JUN-2002; 2002US-0390850P.
PR 29-AUG-2002; 2002US-0407006P.
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Whitney GS, Opitack G, Garulacan L, Ramanathan CS, Mckinnon M;
PI Bennett KL, Barber DE, Cacace A, Tsuchihashi Z;
XX WPI; 2004-090973/09.
XX
XX New nucleic acid molecule encoding a human G-protein coupled receptor,
PT RAI3, useful for preventing, treating or ameliorating chronic obstructive
PT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
PT of COPD.
XX
XX Disclosure; SEQ ID NO 33; 301pp; English.
XX
XX The present sequence is that of DNA encoding a peptide corresponding to
CC amino acids 340-353 of a human G-protein coupled receptor, denoted
CC retinoic acid induced 3 (RAI-3) protein ADI28460. The peptide was
CC identified in studies of RAI-3 phosphorylation in response to cigarette
CC smoke. Primers based on this sequence can be used in a PCR to identify
CC individuals at risk for developing chronic obstructive pulmonary disease
CC (COPD). RAI-3 is tyrosine phosphorylated and/or is associated/complexed
CC with tyrosine phosphorylated proteins only in cells that have been
CC exposed to cigarette smoke. Since RAI-3 is primarily expressed in lung
CC tissue, and since cigarette smoke is a major causative factor of COPD,
CC RAI-3 provides a novel cellular target for identifying modulators, e.g.
CC agonists or antagonists, useful for the treatment and/or prevention of
CC COPD and related disorders such as emphysema and chronic bronchitis. RAI-
CC 3 modulators, e.g. agonists and antagonists, especially antisense
CC compounds, can be used to treat COPD and other disorders an diseases
CC associated with regulation of NF-kB and/or its associated or interacting
CC signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
CC the RAI-3 gene are useful for determining COPD association in
CC individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
CC preventing, treating or ameliorating disorders related to aberrant GPCR
CC signaling or cell cycle regulation, pulmonary disorders, inflammatory
CC lung disorders, COPD, the underlying symptoms of COPD, COPD-related
CC disorders or conditions, autoimmune disorders, disorders related to
CC hyperimmune activity, inflammatory conditions, disorders related to
CC aberrant acute phase responses, hypercongenital conditions, birth
CC defects, necrotic lesions, wounds, organ transplant rejection, renal
CC diseases, ischaemia-reperfusion injury, heart disorders, disorders
CC related to aberrant signal transduction, proliferation disorders,
CC cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
CC ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
CC nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
CC rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
CC pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
CC regulation (all claimed).
XX
XX Sequence 42 BP; 9 A; 8 C; 7 G; 5 T; 0 U; 13 Other;
SQ
Query Match 71.4%; Score 30; DB 12; Length 42;
Best Local Similarity 64.3%; Pred. No. 0.0092;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCGCGAGCCCTTACAAGACTATGAGTAAAG 42
DB 1 GCNCAVCNTGCGCCNTCCNCCNTAARGAYTAGAGTNAAR 42
RESULT 43
ADI28527
ID ADI28527 standard; DNA; 40 BP.
XX
```

XX	AD128527;	
AC		
XX	22-APR-2004 (first entry)	
DT		
XX		
XX	Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA.	
DE		
XX		
KW	Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;	
KW	antiinflammatory; immunosuppressive; cyostatic; cardiant; antiallergic;	
KW	broncholytic; gene therapy; ss.	
XX		
OS	homo sapiens.	
XX		
XX	WO2004001060-A2.	
PN		
PD	31-DEC-2003.	
XX		
XX	20-JUN-2003; 2003WO-US019255.	
PF		
XX		
XX	20-JUN-2002; 2002US-0390850P.	
PR		
PR	29-AUG-2002; 2002US-0407008P.	
XX		
XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	
PA		
PI	Whitney KS, Opitack G, Garulacan L, Ramanathan CS, McKinnon M;	
PI	Bennett GL, Barber LE, Cacace A, Tsuchihashi Z;	
XX		
DR	WPI; 2004-090973/09.	
DR	P-PSDB; AD128458.	
XX		
XX	New nucleic acid molecule encoding a human G-protein coupled receptor,	
PT	RAI3, useful for preventing, treating or ameliorating chronic obstructive	
PT	pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms	
PT	of COPD.	
XX		
PS	Disclosure; SEQ ID NO 30; 30pp; English.	
XX		
CC	The present sequence is that of DNA encoding a peptide corresponding to	
CC	amino acids 340-353 of a human G-protein coupled receptor, denoted	
CC	retinoic acid induced 3 (RAI-3) protein AD128460. The peptide was	
CC	identified in studies of RAI-3 phosphorylation in response to cigarette	
CC	smoke. Primers based on this sequence can be used in a PCR to identify	
CC	individuals at risk for developing chronic obstructive pulmonary disease	
CC	(COPD). RAI-3 is tyrosine phosphorylated and/or is associated/complexed	
CC	with tyrosine phosphorylated proteins only in cells that have been	
CC	exposed to cigarette smoke. Since RAI-3 is primarily expressed in lung	
CC	tissue, and since cigarette smoke is a major causative factor of COPD	
CC	RAI-3 provides a novel cellular target for identifying modulators, e.g.	
CC	agonists or antagonists, useful for the treatment and/or prevention of	
CC	COPD and related disorders such as emphysema and chronic bronchitis. RAI-	
CC	3 modulators, e.g. agonists and antagonists, especially antisense	
CC	compounds, can be used to treat COPD and other disorders an diseases	
CC	associated with regulation of NF-kB and/or its associated or interacting	
CC	signaling molecules. Single nucleotide polymorphisms (SNPs) detected in	
CC	the RAI-3 gene are useful for determining COPD association in	
CC	individuals. RAI-3 nucleic acid molecules and polypeptides are useful for	
CC	preventing, treating or ameliorating disorders related to aberrant GPCR	
CC	signaling or cell cycle regulation, pulmonary disorders, inflammatory	
CC	lung disorders, COPD, the underlying symptoms of COPD, COPD-related	
CC	disorders or conditions, autoimmune disorders, disorders related to	
CC	hyperimmune activity, inflammatory conditions, disorders related to	
CC	aberrant acute phase responses, hypercongenital conditions, birth	
CC	defects, necrotic lesions, wounds, organ transplant rejection, renal	
CC	diseases, ischaemia-reperfusion injury, heart disorders, disorders	
CC	related to aberrant signal transduction, proliferation disorders,	
CC	cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,	
CC	ulcerative colitis, cerebral infarct, myocardial infarct, diabetic	
CC	neuropathy, allergic rhinitis, Crohn's disease, atherosclerosis,	
CC	rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,	
CC	pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,	
CC	colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders	
CC	associated with aberrant cell adhesion, I-CAM function and/or regulation	
CC	E-selectin function and/or regulation, or aberrant NF-kB function and/or	

CC regulation (all claimed).
 XX Sequence 40 BP; 11 A; 12 C; 10 G; 7 T; 0 U; 0 Other;
 SQ

Query Match 66.7%; Score 28; DB 12; Length 40;
 Best Local Similarity 97.5%; Pred. No. 0.065;
 Matches 39; Conservative 0; Mismatches 1; Gaps 1;

Oy 1 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAA 40
 Db |||||
 1 GCCACGCTTGGCGGAGCCCTTAC-AAGACTATGAAGTAA 39

RESULT 44
 ADV92548
 ID ADV92548 standard; DNA; 480 BP.
 XX AC
 XX ADV92548;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 XX Japanese encephalitis virus (JEV) envelope protein (EB) encoding DNA.
 DE
 XX gene; ds; Immune stimulation; virucide; vaccine; envelope protein;
 XX immunostimulant; flavivirus infection.
 KW
 XX Japanese encephalitis virus.
 OS
 XX

Key Location/Qualifiers
 FH 1.480
 FT CDS
 FT /*tag= a
 FT /product= "Japanese encephalitis virus EB protein"
 FT /partial
 FT /note= "No start and stop codons"
 XX
 XX US2004254365-A1.
 PN
 XX
 XX 16-DEC-2004.
 PD
 XX 11-JUN-2003; 2003US-00459155.
 PP
 XX 11-JUN-2003; 2003US-00459155.
 PR
 XX (LEES/) LEE S T.
 PA
 XX
 XX Lee ST;
 PI
 XX
 XX WPI; 2005-030225/03.
 DR P-PSDB; ADV92543.
 DR
 XX

Inducing an immune response against flavivirus, e.g. Japanese encephalitis virus comprises administering a fusion envelope protein of flavivirus or an expression vector containing a nucleic acid encoding the fusion polypeptide.

PS Disclosure; SEQ ID NO 7; 18pp; English.
 XX

The present invention features a method of inducing an immune response in a subject such as human or a non-human animal against a flavivirus. The invention is useful for inducing an immune response in a subject against flavivirus including Dengue virus, tick borne encephalitis virus, West Nile encephalitis virus and Japanese encephalitis virus (JEV). The present sequence is Japanese encephalitis virus (JEV) envelope protein (EB) encoding DNA.

XX
 XX Sequence 480 BP; 118 A; 123 C; 133 G; 106 T; 0 U; 0 Other;
 SQ

Query Match 54.3%; Score 22.8; DB 14; Length 480;
 Best Local Similarity 79.4%; Pred. No. 18;
 Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 4 CACGCTTGGCGGAGCCCTTACAAAGACTATGAAG 37
 Db |||||

Db 330 CACGCTGGCGAAGCCTTTTCAACGACTTTGAAG 363

RESULT 45
ADV92553
ID ADV92553 standard; DNA; 525 BP.
XX
AC ADV92553;
XX
DT 10-MAR-2005 (first entry)
XX
DE Japanese encephalitis virus (JEV) M15EB protein encoding DNA.
XX
KW gene; ds; Immune stimulation; virucide; vaccine; envelope protein;
KW membrane protein; immunostimulant; flavivirus infection.
XX
OS Japanese encephalitis virus.
XX
FH Key Location/Qualifiers
FH CDS 1..525
FT /tag= b
FT /product= "Japanese encephalitis virus M15EB protein"
FT /transl_except= (pos:1..5, aa:Val)
FT /partial
FT /notes= "No start and stop codons"
FT misc_feature 1..45
FT /tag= a
FT /note= "JEV M carboxyl-terminal peptide encoding DNA
FT fragment"
FT /note= "This fragment is stated as signal peptide
FT encoding DNA fragment"
FT misc_feature 46..525
FT /tag= c
FT /note= "JEV EB carboxyl-terminal protein encoding DNA"
XX
XX US2004254365-A1.
XX
PN 16-DEC-2004.
XX
PD 11-JUN-2003; 2003US-00459155.
XX
PR 11-JUN-2003; 2003US-00459155.
XX
PA (LEES/) LEE S T.
XX
XX Lee ST;
XX
DR WPI; 2005-030225/03.
DR P-PSDB; ADV92552.
XX
PT Inducing an immune response against flavivirus, e.g. Japanese
PT encephalitis virus comprises administering a fusion envelope protein of
PT flavivirus or an expression vector containing a nucleic acid encoding the
PT fusion polypeptide.
XX
PS Disclosure; SEQ ID NO 12; 18pp; English.
XX
XX The present invention features a method of inducing an immune response in
CC a subject such as human or a non-human animal against a flavivirus. The
CC invention is useful for inducing an immune response in a subject against
CC flavivirus including Dengue virus, tick borne encephalitis virus, West
CC Nile encephalitis virus and Japanese encephalitis virus (JEV). The
CC present sequence is Japanese encephalitis virus (JEV) M15EB protein
CC encoding DNA. This DNA includes a 45 bp DNA fragment of the carboxyl-
CC terminal part of the JEV membrane (M) protein encoding DNA and a 480 bp
CC DNA fragment EB of the carboxyl-terminal part of the JEV envelope (E)
CC protein encoding DNA.
XX
XX Sequence 525 BP; 123 A; 136 C; 144 G; 122 T; 0 U; 0 Other;
SQ

Query Match 54.3%; Score 22.8; DB 14; Length 525;
Best Local Similarity 79.4%; Pred. No. 19;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 375 CACGCTGGCGAAGCCTTTTCAACGACTTTGAAG 408

RESULT 46
AAA29595
ID AAA29595 standard; cDNA; 1500 BP.
XX
AC AAA29595;
XX
DT 09-AUG-2000 (first entry)
XX
DE Japanese encephalitis virus protein JMSV001 encoding cDNA.
XX
KW Japanese encephalitis virus; infection; immunogen; vaccine; diagnosis;
KW viral particle; ThCAMAR67/93; JMSV001; antiviral; immunomodulatory; ss.
XX
OS Japanese encephalitis virus.
XX
PN WO200020565-A1.
XX
PD 13-APR-2000.
XX
PF 02-JUN-1999; 99WO-JP002931.
XX
PR 05-OCT-1998; 98JP-00319762.
XX
XX (OSAU) UNIV OSAKA.
XX
PI Ishikawa T, Yoshii H, Onishi T, Imagawa T, Ishibashi M;
XX
DR WPI; 2000-303759/26.
DR P-PSDB; AAY82692.
XX
XX Novel inactivated viral particles, useful as vaccines against and in the
XX diagnosis of infection with Japanese encephalitis viruses, is prepared
XX from an infective cell culture of a Japanese encephalitis virus.
XX
PS Disclosure; Page 33-36; 42pp; Japanese.
XX
XX The present invention describes inactivated immunogenic viral particles
CC (I) prepared from an infective cell culture of a Japanese encephalitis
CC virus. (I) are useful for treatment, diagnosis and as vaccines against
CC Japanese encephalitis virus infection. The present sequence encodes a
CC Japanese encephalitis virus protein, designated JMSV001, which is given
CC in the exemplification of the present invention
XX
XX Sequence 1500 BP; 393 A; 362 C; 411 G; 334 T; 0 U; 0 Other;
SQ

Query Match 54.3%; Score 22.8; DB 3; Length 1500;
Best Local Similarity 79.4%; Pred. No. 24;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
DB 1203 CACGCTGGCGAAGCCTTTTCAACGACTTTGAAG 1236

RESULT 47
ADV92550
ID ADV92550 standard; DNA; 1500 BP.
XX
AC ADV92550;
XX
DT 10-MAR-2005 (first entry)
XX
DE Japanese encephalitis virus (JEV) E gene.
XX
KW gene; ds; Immune stimulation; virucide; vaccine; envelope protein;
KW membrane protein; immunostimulant; flavivirus infection.
XX

Query Match 54.3%; Score 22.8; DB 14; Length 525;
Best Local Similarity 79.4%; Pred. No. 19;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;


```

FH key                               Location/Qualifiers
FT sig_peptide 10..78
FT CDS          /*tag= a
FT              /*tag= b
FT /product= "Japanese encephalitis virus prM protein"
FT mat_peptide 355..579
FT              /*tag= c
FT /product= "Japanese encephalitis virus M protein"
FT CDS          580..2079
FT              /*tag= d
FT /product= "Japanese encephalitis virus E protein"
FT JP2004065118-A.
FT 04-MAR-2004.
XX
XX 07-AUG-2002; 2002JP-00229597.
XX
XX 07-AUG-2002; 2002JP-00229597.
XX
XX (HAND-) ZH HANDAI BISEIBUTSUBYO KENKYUKAI.
XX (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.
XX
XX WPI; 2004-209123/20.
XX P-PSDB; ADM97134, ADM97135, ADM97136, ADM97137.
XX
XX Novel Japanese-encephalitis-virus antigen comprising Japanese-
XX encephalitis-virus-like particle consisting of M and E protein, useful as
XX Japanese-encephalitis vaccine.
XX
XX Claim 3; SEQ ID NO 4; 45pp; Japanese.
XX
XX The invention relates to a novel Japanese-encephalitis-virus antigen
XX comprising a Japanese-encephalitis-virus-like particle, which includes
XX the M protein and E protein of a Japanese encephalitis virus, and which
XX does not contain RNA in the particle, and shows haemagglutination
XX activity. The invention further comprises: a transformed cell obtained by
XX providing cDNA encoding the E protein and prM protein prepared from the
XX genome RNA of a Japanese-encephalitis virus, preparing a recombinant DNA
XX by integrating the cDNA in an expression vector which contains a beta-
XX actin promoter and which can be expressed, transforming an animal cell by
XX the recombinant DNA, selecting the transformed cell from a parent cell
XX and making the transformed cell to acclimatize in serum free medium;
XX Japanese-encephalitis-vaccine having immunity for the novel antigen and
XX containing the antigen as an active ingredient; and a diagnostic agent
XX which contains the antigen as an active ingredient. The Japanese-
XX encephalitis-virus antigen is useful for carrying out the preparation of
XX a Japanese-encephalitis vaccine and a diagnostic agent. This
XX polynucleotide sequence represents the Japanese-encephalitis-virus
XX encoding DNA of the invention.
XX
XX Sequence 2088 BP; 555 A; 497 C; 574 G; 462 T; 0 U; 0 Other;
XX
XX Query Match 54.3%; Score 22.8; DB 12; Length 2088;
XX Best Local Similarity 79.4%; Pred. No. 25;
XX Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 4 CACGCTTGCGCGAGCCCTTACAAAGACTATGAAG 37
XX ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1782 CACGCTGGCGAAAGCCTTTTCAACGACTTTGAAG 1815

RESULT 50
AAQ22767
ID AAQ22767 standard; DNA; 4512 BP.
XX
XX AAQ22767;
XX
XX 12-AUG-1992 (first entry)
XX
XX JEV Nakayama strain prM, E, NS1, NS2A, NS2B and C coding regions.
XX

```

```

KW Japanese Encephalitis Virus; vaccinia virus donor; plasmid pDr20; ss.
XX
XX Japanese encephalitis virus.
XX
XX WO9203545-A.
XX
XX 05-MAR-1992.
XX
XX 05-AUG-1991; 91WO-US005816.
XX
XX 15-AUG-1990; 90US-00567960.
XX 06-JUN-1991; 91US-00711429.
XX 13-JUN-1991; 91US-00714687.
XX 17-JUL-1991; 91US-00729800.
XX 05-AUG-1991; 91WO-U0005816.
XX
XX (VIRO-) VIROGENETICS CORP.
XX
XX Paoletti E, Pinc, Pinc, Pincus SE;
XX
XX WPI; 1992-096889/12.
XX
XX Recombinant pox-virus e.g. vaccinia, fowl-pox and canary-pox virus -
XX contg. DNA from flavi-virus e.g. Japanese encephalitis and yellow fever
XX virus, used as vaccine.
XX
XX Example 9; Fig 17; 117pp; English.
XX
XX cDNA was prepared from genomic virion RNA obtained from suspension
XX cultures of C6/36 cells infected with a passage 55 suckling mouse brain
XX stock of the Nagayama strain of JEV. EcoRI linkers were ligated to the
XX cDNA fragments for cloning into pBR322. Recombinant plasmids were
XX transformed into E.coli DH5 cells. Plasmid pC20 contained 81 non-coding
XX nucleotides and the coding regions for C and prM. Sequence AAQ22767 is
XX that of the C coding region of pC20, combined with an updated sequence of
XX the prM, E, NS1, NS2A and NS2B coding regions of the Nagayama strain. The
XX sequence begins at the C protein Met initiation codon. A subfragment of
XX pC20 was cloned into pUC18 to give pDr20. This plasmid was then used in
XX the construction of novel recombinants JEV24, JEV27, JEV33 and JEV34.
XX These were transfected into VP410 infected cells to generate VP825,
XX VP829, VP857 and VP864, respectively
XX
XX Sequence 4512 BP; 1192 A; 1055 C; 1253 G; 1012 T; 0 U; 0 Other;
XX
XX Query Match 54.3%; Score 22.8; DB 2; Length 4512;
XX Best Local Similarity 79.4%; Pred. No. 30;
XX Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 4 CACGCTTGCGCGAGCCCTTACAAAGACTATGAAG 37
XX ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2085 CACGCTGGCGAAAGCCTTTTCAACGACTTTGAAG 2118

RESULT 51
AA80297
ID AA80297 standard; RNA; 4992 BP.
XX
XX AA80297;
XX
XX 25-MAR-2003 (revised)
XX 16-SEP-1990 (first entry)
XX
XX Partial nucleic acid sequence of the Japanese encephalitis virus (JEV)
XX genome encoding C, M, E and NS proteins.
XX
XX Japanese encephalitis virus (JEV); vaccine; flavivirus immunity;
XX C protein; M protein; E protein; NS protein.
XX
XX Japanese encephalitis virus.
XX
XX Key Location/Qualifiers
XX mat_peptide 1..159
XX /*tag= a
XX

```

```
FT /product= "C protein"
FT 160..441
FT /tag= b
FT /product= "pre M protein"
FT 442..666
FT /tag= c
FT /product= "M protein"
FT 667..2166
FT /tag= d
FT /product= "E protein"
FT 2167..3408
FT /tag= e
FT /product= "NS1"
FT 3409..3903
FT /tag= f
FT /product= "ns2a"
FT 3904..4296
FT /tag= g
FT /product= "ns2b"
FT 4297..4992
FT /tag= h
FT /product= "NS3"
XX
XX WO8903032-A.
XX
XX 05-MAY-1988.
XX
XX 21-OCT-1987; 87WO-US002763.
XX
XX 27-OCT-1986; 86US-00923907.
XX
XX (FOUR/) FOURNIER M J.
XX
XX Fournier M, Mason TL, Mcada P, Mason PW;
XX WPI; 1988-133144/19.
XX P-PSDB; AAP80264.
XX
XX Nucleic acid of Japanese encephalitis virus - used for producing
XX polypeptide(s) and in the diagnosis of and immunisation against the
XX virus.
XX
XX Claim 4; Fig 5/1-5/5; 52pp; English.
XX
XX Substantially purified nucleic acid comprising at least a 10 bp sequence
XX of DNA corresponding to the nucleic acid sequence of JEV but not to the
XX nucleic acid sequence of yellow fever virus wherein the 10 bp sequence is
XX included within the sequence shown in AAN80297. Pref. the nucleic acid
XX sequences do not correspond to those found in yellow fever virus, West
XX Nile virus, Murray valley fever virus, dengue virus or St Louis
XX encephalitis virus. The nucleic acid may be obtd. from the virus and
XX inserted into a vector. Specific probes can be derived from this nucleic
XX acid. Also claimed are a pure polypeptide synthesized from the purified
XX nucleic acid and antibodies produced to the polypeptide. The polypeptides
XX and nucleic acid prods. are used in the diagnosis of and immunisation
XX against JEV. Inoculation of animals with synthetic viral protein
XX immunogens or with active vectors encoding such immunogens, elicits
XX protective antiviral antibodies. The vaccination induces immunity to the
XX disease caused by other flaviviruses. (Updated on 25-MAR-2003 to correct
XX PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 4992 BP; 1332 A; 1166 C; 1390 G; 3 T; 1100 U; 1 Other;
XX
XX Query Match 54.3%; Score 22.8; DB 1; Length 4992;
XX Best Local Similarity 64.7%; Pred. No. 31;
XX Matches 22; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
XX
XX 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
XX ||||| ||| ||||| : ||||| : |||||
XX 1869 CACGUGGCGAAAGCCUUCUACGACGACUUGAAG 1902
XX
XX RESULT 52
XX
XX Query Match 54.3%; Score 22.8; DB 14; Length 7486;
XX Best Local Similarity 79.4%; Pred. No. 34;
XX Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
XX ||||| ||| ||||| ||||| ||||| |||||
XX 2710 CACGCTGGCGAAAGCCCTTTCAACGACTTTGAAG 2743
XX
XX RESULT 53
XX ABK62169
XX ID ABK62169 standard; cDNA; 584 BP.
XX
XX AC ABK62169;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Rat sequence differentially expressed in response to a hepatotoxin #76.
XX
XX KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
XX OS Rattus norvegicus.
XX
XX PN WO200210453-A2.
XX
XX 07-FEB-2002.
XX
ADW74114
ID ADW74114 standard; DNA; 7486 BP.
XX
XX AC ADW74114;
XX
XX DT 07-APR-2005 (first entry)
XX
XX DE pcJEME DNA (Japanese encephalitis virus prM/E gene in pcDNA3 vector).
XX
XX KW nucleic acid vaccine; gene transfer; vector; ds.
XX
XX OS Japanese encephalitis virus strain Nakayama.
XX Synthetic.
XX
XX PN JP2005015355-A.
XX
XX PD 20-JAN-2005.
XX
XX PF 24-JUN-2003; 2003JP-00179620.
XX
XX PR 24-JUN-2003; 2003JP-00179620.
XX
XX PA (UYKO-) UNIV KOBE.
XX
XX DR WPI; 2005-105264/12.
XX
XX Increasing amount of antigen produced from DNA vaccine, by administering
XX DNA vaccine using needle-less syringe, which produce increased amount of
XX antigen than produced by administering DNA vaccine using usual syringe
XX with needle.
XX
XX Example; SEQ ID NO 1; 22pp; Japanese.
XX
XX The invention relates to a novel method for increasing the amount of
XX antigen produced from a DNA vaccine. The method comprises administering
XX the DNA vaccine using a needle-less syringe, whereby using such a syringe
XX the amount of antigen produced is increased when compared to the amount
XX of antigen produced by administering DNA vaccine using a syringe with a
XX needle. The DNA vaccine is administered by an intramuscular route. The
XX method of the invention may be useful for increasing the amount of
XX antigen produced from a DNA vaccine. The current sequence is that of the
XX pcJEME DNA of the invention which comprises the Japanese encephalitis
XX virus strain Nakayama prM/E gene in the pcDNA3 vector.
XX
XX Sequence 7486 BP; 1791 A; 1909 C; 1948 G; 1838 T; 0 U; 0 Other;
XX
XX Query Match 54.3%; Score 22.8; DB 14; Length 7486;
XX Best Local Similarity 79.4%; Pred. No. 34;
XX Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 4 CACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
XX ||||| ||| ||||| ||||| ||||| |||||
XX 2710 CACGCTGGCGAAAGCCCTTTCAACGACTTTGAAG 2743
XX
XX RESULT 53
XX ABK62169
XX ID ABK62169 standard; cDNA; 584 BP.
XX
XX AC ABK62169;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Rat sequence differentially expressed in response to a hepatotoxin #76.
XX
XX KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
XX OS Rattus norvegicus.
XX
XX PN WO200210453-A2.
XX
XX 07-FEB-2002.
```

```
XX 30-JUL-2001; 2001WO-US023872.
PP
XX 31-JUL-2000; 2000US-0222040P.
XX 02-NOV-2000; 2000US-0244880P.
PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
PI WPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or cells
PT exposed to the toxin and comparing these to gene expression in unexposed
PT tissues or cells.
XX
XX Claim 1; SEQ ID NO 76; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic effect
CC of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression. The
CC method can also be used to identify an agent which modulates the toxic
CC response and predict cellular pathways that a compound modulates in a
CC cell. The methods utilise a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent
XX
XX Sequence 584 BP; 207 A; 108 C; 91 G; 178 T; 0 U; 0 Other;
SQ
    Query Match      52.9%; Score 22.2; DB 6; Length 584;
    Best Local Similarity 88.9%; Pred. No. 35;
    Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAAGTAA 41
Db 373 GAGCCCTTACAAAGACTATGAAGTAA 399
    |||||
RESULT 54
ADB55155
ID ADB55155 standard; DNA; 584 BP.
XX
XX ADB55155;
AC
```

```
XX 04-DEC-2003 (first entry)
DT
XX Toxicity-related gene, SEQ ID 181.
XX
XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; ds.
XX
XX Unidentified.
OS
XX WO2003064624-A2.
PN
XX 07-AUG-2003.
PD
XX
XX 31-JAN-2003; 2003WO-US003194.
PF
XX 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
PI WPI; 2003-689530/65.
XX
XX Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
PT the compound.
XX
XX Claim 1; SEQ ID NO 181; 1156pp; English.
XX
XX The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 584 BP; 207 A; 108 C; 91 G; 178 T; 0 U; 0 Other;
SQ
    Query Match      52.9%; Score 22.2; DB 10; Length 584;
    Best Local Similarity 88.9%; Pred. No. 35;
    Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAAGTAA 41
Db 373 GAGCCCTTACAAAGACTATGAAGTAA 399
    |||||
RESULT 55
ADB49676
ID ADB49676 standard; DNA; 584 BP.
XX
XX ADB49676;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:218.
DE
XX
```



```
PN US2002137081-A1.
XX
XX
XX 26-SEP-2002.
XX
XX 08-JAN-2002; 2002US-00044090.
XX
XX 28-JUL-2000; 2000US-0222469P.
XX
XX 08-JAN-2001; 2001US-0260483P.
XX
XX (BAND/) BANDMAN O.
XX
XX Bandman O;
XX
XX WPI; 2003-110597/10.
XX
XX Combination for diagnosing, staging, treating, or monitoring the
XX progression of treatment of a vascular disease, e.g. atherosclerosis,
XX PT comprises several cDNAs that are differentially expressed in activated
XX PT vascular tissue.
XX
XX Claim 1; Page; 18pp; English.
XX
XX This invention relates to a combination comprising several cDNAs that are
XX differentially expressed in activated vascular tissue. The invention also
XX CC discloses a high throughput method for detecting differentially expressed
XX CC cDNAs in a sample. The cDNAs of the invention may have
XX CC antiarteriosclerotic; cytostatic; cardiac; hypotensive; antidiabetic;
XX CC gynaecological; vasotropic and cerebroprotective activities and may be
XX CC used in gene therapy. The cDNAs of the invention may be used in a high-
XX CC throughput methods for detecting differential expression of one or more
XX CC cDNAs in a sample, or screening several molecules or compounds to
XX CC identify a molecule or compound that specifically binds a cDNA of the
XX CC invention. A protein encoded by the cDNA may be used to screen several
XX CC molecules or compounds to identify a ligand that specifically binds to
XX CC the protein, or to produce or purify an antibody to the protein that can
XX CC be used to detect a protein in a sample or purify a natural or
XX CC recombinant protein from a sample. The nucleotides may be useful for
XX CC diagnosing, staging, treating, or monitoring the progression of treatment
XX CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
XX CC disease, hypertension, diabetes, pre-eclampsia, ischaemia- reperfusion
XX CC injury, retinosis, or stroke. The cDNAs can also be used for large-scale
XX CC genetic or gene expression analysis of several new nucleic acid
XX CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
XX CC diagnosing pre-pathologic disorders, and chronic or acute disease
XX CC associated with abnormalities in the expression, amount or distribution
XX CC of the protein. The present sequence represents a cDNA of the invention
XX CC that is differentially expressed in activated vascular tissue. Note: The
XX CC sequence data for this patent did not form part of the specification, but
XX CC was obtained in electronic format directly from USPTO at
XX CC http.seqdata.uspto.gov/sequence.html?DocID=20020137081
XX
XX Sequence 4826 BP; 1161 A; 1190 C; 1241 G; 1234 T; 0 U; 0 Other;
XX
XX Query Match 51.9%; Score 21.8; DB 8; Length 4826;
XX Best Local Similarity 70.7%; Pred. No. 82;
XX Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
XX
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
DB 4091 GCGAAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAA 4051

RESULT 58
ADJ56498/c
ID ADJ56498 standard; cDNA; 4826 BP.
XX
XX ADJ56498;
XX
XX 06-MAY-2004 (first entry)
XX
XX Human cDNA differentially expressed in MYCN activated cells SeqID 304.
XX
XX human; differential expression; transactivator; proto-oncogene;
XX
```

```
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
XX Homo sapiens.
XX
XX US2003119009-A1.
XX
XX 26-JUN-2003.
XX
XX 25-FEB-2002; 2002US-00084817.
XX
XX 23-FEB-2001; 2001US-0270784P.
XX
XX (STUA/) STUART S G.
XX (NUCH/) NUCHTERN J G.
XX (PLON/) PLON S E.
XX (SHOH/) SHOHET J M.
XX
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
XX WPI; 2003-635698/60.
XX
XX New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.
XX
XX Claim 1; SEQ ID NO 304; 27pp; English.
XX
XX This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX CC polynucleotide sequences that exhibit differential expression patterns in
XX CC cells activated by the transactivator MYCN, where MYCN is a proto-
XX CC oncogene that is amplified in neuroblastoma cells and is common in small
XX CC cell lung cancers. The present invention describes these cDNA molecules
XX CC as useful for in hybridisation assays to detect expression of nucleic
XX CC acids (or complementary nucleic acids) in a present in a given sample, as
XX CC well as for screening assays by identifying molecules or compounds that
XX CC specifically bind the cDNA as a ligand and modulate function or activity.
XX CC Accordingly, these compositions exhibit cytostatic activity and can also
XX CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX CC that is differentially expressed in MYCN activated cells, given in an
XX CC exemplification of the invention. NOTE: This sequence does not appear in
XX CC the printed specification but has been obtained in electronic format from
XX CC the US Patent Office at
XX CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX
XX Sequence 4826 BP; 1161 A; 1190 C; 1241 G; 1234 T; 0 U; 0 Other;
XX
XX Query Match 51.9%; Score 21.8; DB 10; Length 4826;
XX Best Local Similarity 70.7%; Pred. No. 82;
XX Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
XX
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
DB 4091 GCGAAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAA 4051

RESULT 59
ABK42428
ID ABK42428 standard; DNA; 17849 BP.
XX
XX ABK42428;
XX
XX 21-MAY-2002 (first entry)
XX
XX Genomic sequence #327 encoding novel human connective tissue polypeptide.
XX Human; connective tissue related disorder; cancer; gene therapy;
XX cytostatic; gene; ds.
XX
XX Homo sapiens.
XX
XX
```

PN WO200155343-A1.
 XX 02-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US001322.
 PF
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0255671P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;


```
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 10139; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX
XX Sequence 17849 BP; 3418 A; 5243 C; 5427 G; 3761 T; 0 U; 0 Other;
SQ
Query Match 51.9%; Score 21.8; DB 4; Length 17849;
Best Local Similarity 70.7%; Pred. No. 1.le+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 2 CCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 42
DB 13871 CCCCCACTTCTAAGAGGCTTTATAAAACTTTCACGTAAG 13911
. RESULT 61
```

```
ADB60584
ID ADB60584 standard; DNA; 17849 BP.
XX
AC ADB60584;
XX
DT 04-DEC-2003 (first entry)
XX
DE Connective tissue related genomic DNA #327.
XX
XX cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
XX antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;
XX antiinflammatory; antiallergic; antiasthmatic; dermatological;
XX nephrotoxic; virucide; fungicide; antibacterial; antiparasitic;
XX gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
XX systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
XX cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;
XX Alzheimer's disease; Parkinson's disease; cardiovascular disease;
XX atherosclerosis; myocarditis; cardiopulmonary bypass complication;
XX autoimmune disease; multiple sclerosis; allergic reaction; asthma;
XX rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
XX gastrointestinal disorder; inflammatory bowel disease;
XX organ transplant rejection; immune system disorder; Bruton's disease;
XX X-linked lymphoproliferative syndrome;
XX B-cell lymphoproliferative disorder; HIV; AIDS; infection;
XX chromosome identification; chromosome mapping; ds.
XX connective tissue related polynucleotide; gene; ds.
XX
XX Homo sapiens.
XX
XX US2003054375-A1.
XX
XX 20-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00092154.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 22-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
```


DE Genomic sequence #325 encoding novel human connective tissue polypeptide.
XX
KW Human; connective tissue related disorder; cancer; gene therapy;
KW cytostatic; gene; ds.
OS Homo sapiens.
XX WO20015343-A1.
PN XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001322.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232977P.
PR 14-SEP-2000; 2000US-0232988P.
PR 14-SEP-2000; 2000US-0232999P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234221P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 13-OCT-2000; 2000US-023937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

```
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-565190/63.
XX Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder such
PT as cancer or rheumatoid arthritis.
XX Disclosure; SEQ ID NO 1313; 673pp; English.
XX The present invention relates to the isolation of novel human connective
CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
CC (cDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful in
CC gene therapy. ABK42102-ABK43116 represent genomic sequences encoding the
CC novel human connective tissue related polypeptides. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 17862 BP; 3420 A; 5246 C; 5430 G; 3766 T; 0 U; 0 Other;
Query Match 51.9%; Score 21.8; DB 4; Length 17862;
Best Local Similarity 70.7%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 2 CCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 13885 CCCCACCTTCTAAGAGGCTTTATAAAACTTTCAAGTAAAG 13925
RESULT 63
AAL07450
ID AAL07450 standard; DNA; 17862 BP.
XX
AC AAL07450;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 10138.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US0011339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235836P.
PR 27-SEP-2000; 2000US-0235837P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
```


PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234957P.
PR 25-SEP-2000; 2000US-0234988P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 17-JAN-2001; 2001US-00764847.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-634869/60.

DR P-PSDB; ADB59454.

XX New connective tissue-related polypeptides and polynucleotides, useful
PT for treating, preventing and/or prognosing e.g. disorders of connective
PT tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or
PT neoplasias.

XX Disclosure; SEQ ID NO 1313; 248pp; English.

XX The invention describes an isolated nucleic acid molecule (I), which
XX comprises a sequence that is at least 95 % identical to a connective
XX tissue-related polynucleotide encoding connective tissue antigens (CTA).
XX The polypeptide or polynucleotide is useful for preventing, treating, or
XX ameliorating medical conditions in a mammal. The connective tissue
XX polypeptides, polynucleotides and antibodies are particularly useful for
XX treating, preventing and/or prognosing disorders of connective tissues
XX (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
XX scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
XX neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.
XX Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
XX (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass
XX complications), autoimmune diseases (e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.

Query Match 51.9%; Score 21.8; DB 9; Length 17862;
Best Local Similarity 70.7%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Oy 2 CCCACGCTGGCCGAGCCCTTACAAAGACTATCAAGTAAAG 42

Db 13885 CCCCCCTTAAAGAGGCTTTATAAAACCTTCAAGTAAAG 13925

RESULT 65

```
ABLI17076
ID ABLI17076 standard; DNA; 21748 BP.
XX
AC ABLI17076;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 2701.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 2701; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABBS57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 21748 BP; 6384 A; 4237 C; 4321 G; 6806 T; 0 U; 0 Other;

Query Match 51.9%; Score 21.8; DB 4; Length 21748;
Best Local Similarity 78.8%; Pred. No. 1.1e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTGGCGGAGCCCTTACAAAGACTATGAACTAAA 41
|||||
DB 8163 TTGGGTAGCCGTTCAATGACTTTGAAATAAA 8195

RESULT 66
AED85840/c
ID AED85840 standard; DNA; 363 BP.
XX
AC AED85840;
XX
DT 12-JAN-2006 (first entry)
XX
DE Human CD38 antibody heavy chain variable region 3077_VH1B DNA.
XX
KW immunosuppressive; hematological; cytostatic; antiinflammatory;
KW antirheumatic; antiarthritic; antibody; CD38; hematological disease;
KW inflammation; heavy chain variable region; ds; coding sequence.
XX
OS Homo sapiens.
XX

ABLI17076
Key Location/Qualifiers
CDS 1..363
FT /*tag= a
FT /partial
FT /product= "3077 VH1B"
FT /note= "No start or stop codon given"
XX
PN WO2005103083-A2.
XX
XX
PD 03-NOV-2005.
XX
PF 07-FEB-2005; 2005WO-IB002476.
XX
PR 06-FEB-2004; 2004US-0541911P.
PR 26-FEB-2004; 2004US-0547584P.
PR 18-MAR-2004; 2004US-0553948P.
PR 06-AUG-2004; 2004US-0599014P.
XX
PA (MORP-) MORPHOSYS AG.
XX
PI Tesar M, Jager U;
XX
DR WPI; 2005-734713/75.
DR P-PSDB; AED85844.
XX
XX New isolated human or humanized antibody or its functional fragment
PT comprising an antigen-binding region that is specific for an epitope of
PT CD38, useful for treating hematological or inflammatory disorders.
XX
PS Claim 62; SEQ ID NO 1; 98pp; English.
XX
CC The invention relates to an isolated human or humanized antibody or its
CC functional fragment comprising an antigen-binding region that is specific
CC for an epitope of CD38. The antibody or its functional fragment is useful
CC for treating a disorder or condition associated with the undesired
CC presence of CD38+ cells, e.g. hematological disease, such as multiple
CC myeloma, chronic lymphocytic leukemia, chronic myelogenous leukemia,
CC acute myelogenous leukemia, and acute lymphocytic leukemia; or an
CC inflammatory disease such as rheumatoid arthritis and systemic lupus
CC erythematosus. The present sequence represents the human CD38 antibody
CC heavy chain variable region 3077_VH1B DNA.
XX
SQ Sequence 363 BP; 85 A; 81 C; 107 G; 90 T; 0 U; 0 Other;

Query Match 51.4%; Score 21.6; DB 14; Length 363;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 7 GCTTGGCGGAGCCCTTACAAAGACTATGAACTAAAAG 42
|||||
DB 119 GCTTGGCGGAGCCCAATTATAGATAAGAAAGTAAAG 84

RESULT 67
AED85870/c
ID AED85870 standard; DNA; 363 BP.
XX
AC AED85870;
XX
DT 12-JAN-2006 (first entry)
XX
DE Human CD38 antibody heavy chain variable region 3077_VH1B DNA.
XX
KW immunosuppressive; hematological; cytostatic; antiinflammatory;
KW antirheumatic; antiarthritic; antibody; CD38; hematological disease;
KW inflammation; heavy chain variable region; ds; coding sequence.
XX
OS Homo sapiens.
XX
PN WO2005103083-A2.
XX
PD 03-NOV-2005.
XX
```

```
PF 07-FEB-2005; 2005WO-IB002476.
XX
XX 06-FEB-2004; 2004US-0541911P.
XX 26-FEB-2004; 2004US-0547584P.
XX 18-MAR-2004; 2004US-0553948P.
XX 06-AUG-2004; 2004US-0559014P.
XX (MORP-) MORPHOSYS AG.
XX
XX Tesar M, Jager U;
XX
XX WPI; 2005-734713/75.
XX
XX New isolated human or humanized antibody or its functional fragment
XX comprising an antigen-binding region that is specific for an epitope of
XX CD38, useful for treating hematological or inflammatory disorders.
XX
XX Disclosure; SEQ ID NO 31; 98pp; English.
XX
XX The invention relates to an isolated human or humanized antibody or its
XX functional fragment comprising an antigen-binding region that is specific
XX for an epitope of CD38. The antibody or its functional fragment is useful
XX for treating a disorder or condition associated with the undesired
XX presence of CD38+ cells, e.g. hematological disease, such as multiple
XX myeloma, chronic lymphocytic leukemia, chronic myelogenous leukemia,
XX acute myelogenous leukemia, and acute lymphocytic leukemia; or an
XX inflammatory disease such as rheumatoid arthritis and systemic lupus
XX erythematosus. The present sequence represents the human CD38 antibody
XX heavy chain variable region 3077_i_VH1B DNA.
XX
XX Sequence 363 BP; 87 A; 80 C; 105 G; 91 T; 0 U; 0 Other;
XX
XX Query Match 51.4%; Score 21.6; DB 14; Length 363;
XX Best Local Similarity 75.0%; Pred. No. 56;
XX Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX
XX 7 GCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
XX 119 GCTTGGGGGCCCAATTATAGATTAAGAGTAAAG 84
XX
XX RESULT 68
XX AAC16864/c
XX ID AAC16864 standard; cDNA; 452 BP.
XX
XX AAC16864;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 20939.
XX
XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
```

```
XX Claim 1; SEQ ID NO 20939; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors
XX
XX Sequence 452 BP; 127 A; 129 C; 109 G; 84 T; 0 U; 3 Other;
XX
XX Query Match 51.4%; Score 21.6; DB 3; Length 452;
XX Best Local Similarity 75.0%; Pred. No. 59;
XX Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX
XX 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAA 36
XX 400 GTCCAGGCTTGTCACGCCCTTAGAAGGGCTAGCAA 365
XX
XX RESULT 69
XX AAD42934/c
XX ID AAD42934 standard; DNA; 70000 BP.
XX
XX AAD42934;
XX
XX 15-NOV-2002 (first entry)
XX
XX Human phospholipase A2 (PLA2), group VI (Ca2+-independent) gene.
XX
XX Human; antisense; phospholipase A2; infection; inflammation; tumour;
XX antisense therapy; PLA2; ds.
XX
XX Homo sapiens.
XX
XX US6410325-B1.
XX
XX 25-JUN-2002.
XX
XX 09-MAY-2001; 2001US-00851896.
XX
XX 09-MAY-2001; 2001US-00851896.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Bennett CF, Freier SM, Watt AT;
XX
XX WPI; 2002-616513/66.
XX
XX Novel antisense compounds useful for inhibiting gene expression of human
XX phospholipase A2, group VI and for treating diseases associated with
XX expression of phospholipase A2, group VI.
XX
XX Example 13; Col 49-108; 72pp; English.
XX
XX The present invention relates to novel antisense compounds which inhibit
XX the expression of phospholipase A2 (PLA2), group VI (Ca2+-independent).
XX The invention is useful for inhibiting the expression of PLA2, group VI
XX (Ca2+-independent) in human cells or tissues and for treating an animal,
XX particularly a human suspected of having or being prone to a disease or
XX condition associated with expression of human PLA2, group VI (Ca2+-
XX independent). It is useful for diagnostics, therapeutics and as research
XX reagent, e.g. prophylactically to prevent or delay infection, tumour
XX formation or inflammation. The present DNA sequence is human PLA2, group
```

```
CC VI (Ca2+-independent) gene
XX
SQ Sequence 70000 BP; 15386 A; 18193 C; 17857 G; 18563 T; 0 U; 1 Other;
  Query Match 51.4%; Score 21.6; DB 6; Length 70000;
  Best Local Similarity 75.0%; Pred. No. 1.8e+02;
  Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAGCCCTTACAAAGACTATGAA 36
  ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42352 GTCGAGCTTGTCACGCCCTTAGAAGGCTAGCAA 42317

RESULT 70
AAC28686
ID AAC28686 standard; cDNA; 407 BP.
XX
AC AAC28686;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 32761.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW Gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 32761; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 407 BP; 124 A; 78 C; 81 G; 123 T; 0 U; 1 Other;
  Query Match 51.0%; Score 21.4; DB 3; Length 407;
  Best Local Similarity 80.6%; Pred. No. 70;
  Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 11 GGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
  ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 GGCATGCCCTTCACATGACTATGAAGTAAA 40

RESULT 71
ABL03149
ID ABL03149 standard; cDNA; 2356 BP.
XX
AC ABL03149;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 3929.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB59046.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 3929; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2356 BP; 650 A; 606 C; 701 G; 399 T; 0 U; 0 Other;
  Query Match 51.0%; Score 21.4; DB 4; Length 2356;
  Best Local Similarity 71.8%; Pred. No. 1e+02;
  Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAGCCCTTACAAAGACTATGAAGTA 39
  ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1212 GCCCAAGATCGAGAGCCATTATCAAGATATCAAGGA 1250

RESULT 72
ADW52673
ID ADW52673 standard; DNA; 882 BP.
XX
AC ADW52673;
XX
DT 21-APR-2005 (first entry)
XX
DE Japanese encephalitis virus outer coat protein coding seq - SEQ ID 2.
XX
KW vaccine; Japanese encephalitis virus infection; outer coat protein; gene;
KW ds.
```



```
PN US2003233675-A1.
XX
XX 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 36568; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or by
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polynucleotide used in
XX the scope of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1555 BP; 426 A; 371 C; 422 G; 336 T; 0 U; 0 Other;
XX
XX Query Match 50.5%; Score 21.2; DB 13; Length 1555;
XX Best Local Similarity 76.5%; Pred. No. 1.1e+02;
XX Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
QY 2 CCCACGCTTGGCGGAGCCCTTACAAAGACTATGA 35
DB 299 CCAACATTGGCGGTCACCTTACCGTGAATGA 266
XX
RESULT 75
AD170879
ID AD170879 standard; DNA; 527 BP.
XX
XX AD170879;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human ovarian cancer DNA marker #3621.
XX
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
XX Homo sapiens.
XX
XX OS
```

```
XX WO200170979-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009126.
XX
XX 21-MAR-2000; 2000US-0191031P.
XX 25-MAY-2000; 2000US-0207124P.
XX 15-JUN-2000; 2000US-0211940P.
XX 07-JUL-2000; 2000US-0216820P.
XX 25-JUL-2000; 2000US-0220661P.
XX 21-DEC-2000; 2000US-0257672P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lee J, Lillie J;
XX
XX WPI; 2001-611502/70.
XX
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
XX cancer cells as compared to their normal non-cancerous ovarian cells are
XX used to characterize stage, grade, histological type of ovarian cancer.
XX
XX Disclosure; SEQ ID NO 3621; 106pp; English.
XX
XX The invention relates to nucleic acid markers which are overexpressed in
XX ovarian cancer cells as compared to their expression in normal (i.e. non-
XX cancerous) ovarian cells. The invention also relates to polypeptides
XX encoded by the markers, antibodies that selectively bind to the
XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk
XX of developing ovarian cancer involving inhibiting expression of a gene
XX corresponding to a marker of the invention and a method of treating a
XX patient afflicted with ovarian cancer comprising providing to cells of
XX the patient an antisense oligonucleotide complementary to a marker of the
XX invention. The markers are useful for assessing if a patient is afflicted
XX with ovarian cancer, which involves comparing the level of expression of
XX a marker in a patient sample and a normal level of expression of the
XX marker in a control non-ovarian cancer sample. A difference between the
XX expression levels indicates ovarian cancer. The level of expression of a
XX marker corresponds to a secreted protein or to a transcribed
XX polynucleotide or its portion. The level of expression of the marker is
XX assessed by detecting the presence in the sample, a protein or protein
XX fragment corresponding to the marker. The presence of protein or protein
XX fragment is detected using an antibody that specifically binds with the
XX protein or protein fragment. Alternatively, the level of expression of
XX the marker is assessed by detecting the presence of a transcribed
XX polynucleotide which anneals with the marker or anneals with a portion of
XX the polynucleotide comprising the marker under stringent conditions. The
XX marker is also used for monitoring the progression of ovarian cancer in a
XX patient which involves detecting expression of the marker in a patient
XX sample at a first point in time, repeating the method at a subsequent
XX time and comparing the level of expression. The method is carried out
XX using an ovarian tissue sample. A composition comprising a marker,
XX polypeptide or antibody of the invention is used to treat ovarian cancer.
XX This sequence represents a human ovarian cancer DNA marker of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 527 BP; 187 A; 102 C; 116 G; 120 T; 0 U; 2 Other;
XX
XX Query Match 50.0%; Score 21; DB 5; Length 527;
XX Best Local Similarity 73.0%; Pred. No. 1.1e+02;
XX Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 5 AGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAA 41
DB 458 AAGTTTCACAGAGTCCTTCAAAGACTTTGAAGAAAA 494
XX
RESULT 76
AD177193
```


ID ADI77193 standard; DNA; 527 BP.
XX AC ADI77193;
XX DT 20-MAY-2004 (first entry)
XX DE Human ovarian cancer DNA marker #9935.
XX KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX OS Homo sapiens.
XX PN WO200170979-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009126.
XX PR 21-MAR-2000; 2000US-0191031P.
XX PR 25-MAY-2000; 2000US-0207124P.
XX PR 15-JUN-2000; 2000US-0211940P.
XX PR 07-JUL-2000; 2000US-0216820P.
XX PR 25-JUL-2000; 2000US-0220661P.
XX PR 21-DEC-2000; 2000US-0257672P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Lee J, Lillie J;
XX DR WPI; 2001-611502/70.
XX PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
XX PS Disclosure; SEQ ID NO 9935; 106pp; English.
XX CC The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 527 BP; 187 A; 102 C; 116 G; 120 T; 0 U; 2 Other;

Query Match 50.0%; Score 21; DB 5; Length 527;
Best Local Similarity 73.0%; Pred. No. 1.1e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 5 ACCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 41
Db 458 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAA 494
RESULT 77
ADI73027
ID ADI73027 standard; DNA; 559 BP.
XX AC ADI73027;
XX DT 20-MAY-2004 (first entry)
XX DE Human ovarian cancer DNA marker #5769.
XX KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX OS Homo sapiens.
XX PN WO200170979-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009126.
XX PR 21-MAR-2000; 2000US-0191031P.
XX PR 25-MAY-2000; 2000US-0207124P.
XX PR 15-JUN-2000; 2000US-0211940P.
XX PR 07-JUL-2000; 2000US-0216820P.
XX PR 25-JUL-2000; 2000US-0220661P.
XX PR 21-DEC-2000; 2000US-0257672P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Lee J, Lillie J;
XX DR WPI; 2001-611502/70.
XX PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
XX PS Disclosure; SEQ ID NO 5769; 106pp; English.
XX CC The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 559 BP; 198 A; 109 C; 127 G; 124 T; 0 U; 1 Other;

Query Match 50.0%; Score 21; DB 5; Length 559;
Best Local Similarity 73.0%; Pred. No. 1.1e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAA 41
DB 458 AAGTTTCACAGAGTCTCTTCAAAAGACTCTGAAGAAAA 494

RESULT 78
ADL38162

ID ADL38162 standard; DNA; 559 BP.

XX AC ADL38162;

XX DT 20-MAY-2004 (first entry)

XX DE Human ovarian cancer DNA marker #12052.

XX KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX OS Homo sapiens.

XX PN WO200170979-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009126.

XX PR 21-MAR-2000; 2000US-0191031P.

XX PR 25-MAY-2000; 2000US-0207124P.

XX PR 15-JUN-2000; 2000US-0211940P.

XX PR 07-JUL-2000; 2000US-0216820P.

XX PR 25-JUL-2000; 2000US-0220661P.

XX PR 21-DEC-2000; 2000US-0257672P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lee J, Lillie J;
XX WPI; 2001-611502/70.

XX DR Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
XX cancer cells as compared to their normal non-cancerous ovarian cells are
XX used to characterize stage, grade, histological type of ovarian cancer.

XX PS Disclosure; SEQ ID NO 12052; 106pp; English.

XX CC The invention relates to nucleic acid markers which are overexpressed in
XX ovarian cancer cells as compared to their expression in normal (i.e. non-
XX cancerous) ovarian cells. The invention also relates to polypeptides
XX encoded by the markers, antibodies that selectively bind to the
XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk
XX of developing ovarian cancer involving inhibiting expression of a gene
XX corresponding to a marker of the invention and a method of treating a
XX patient afflicted with ovarian cancer comprising providing to cells of
XX the patient an antisense oligonucleotide complementary to a marker of the
XX invention. The markers are useful for assessing if a patient is afflicted
XX with ovarian cancer, which involves comparing the level of expression of
XX a marker in a patient sample and a normal level of expression of the
XX marker in a control non-ovarian cancer sample. A difference between the
XX expression levels indicates ovarian cancer. The level of expression of a

CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.

XX Sequence 559 BP; 198 A; 109 C; 127 G; 124 T; 0 U; 1 Other;

Query Match 50.0%; Score 21; DB 5; Length 559;
Best Local Similarity 73.0%; Pred. No. 1.1e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAA 41
DB 458 AAGTTTCACAGAGTCTCTTCAAAAGACTCTGAAGAAAA 494

RESULT 79

ADL36609

ID ADL36609 standard; DNA; 579 BP.

XX AC ADL36609;

XX DT 20-MAY-2004 (first entry)

XX DE Human ovarian cancer DNA marker #10499.

XX KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX OS Homo sapiens.

XX PN WO200170979-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009126.

XX PR 21-MAR-2000; 2000US-0191031P.

XX PR 25-MAY-2000; 2000US-0207124P.

XX PR 15-JUN-2000; 2000US-0211940P.

XX PR 07-JUL-2000; 2000US-0216820P.

XX PR 25-JUL-2000; 2000US-0220661P.

XX PR 21-DEC-2000; 2000US-0257672P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lee J, Lillie J;
XX WPI; 2001-611502/70.

XX PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
XX cancer cells as compared to their normal non-cancerous ovarian cells are
XX used to characterize stage, grade, histological type of ovarian cancer.

XX PS Disclosure; SEQ ID NO 10499; 106pp; English.

XX CC The invention relates to nucleic acid markers which are overexpressed in
XX ovarian cancer cells as compared to their expression in normal (i.e. non-
XX cancerous) ovarian cells. The invention also relates to polypeptides
XX encoded by the markers, antibodies that selectively bind to the
XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk

of detecting ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct sequences.

PT	Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
PT	Disclosure; SEQ ID NO 4193; 106pp; English.
XX	The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences .
XX	Sequence 579 BP; 205 A; 114 C; 134 G; 125 T; 0 U; 1 Other;
XX	Query Match 50.0%; Score 21; DB 5; Length 579;
XX	Best Local Similarity 73.0%; Pred. No. 1.1e+02;
XX	Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy	5 ACGCTTGGCCGAGCCCTTCAAAAGACTATCAAGTAAA 41
Db	459 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAA 495
ADL44546	
ID	ADL44546 standard; DNA; 625 BP.
AC	ADL44546;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Human ovarian cancer DNA marker #18436.
XX	Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX	Homo sapiens.
OS	
XX	WO200170379-A2.
XX	
PD	27-SEP-2001.
XX	
PF	21-MAR-2001; 2001WO-US0009126.
XX	
PR	21-MAR-2000; 2000US-0191031P.
PR	25-MAY-2000; 2000US-0207124P.

ID AD030325 standard; cDNA; 1068 BP.
AC AD030325;
XX
DT 29-JUL-2004 (first entry)
XX Mouse GPCR RAI3 polynucleotide, SEQ ID NO:1428.
DE
XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cytotatic; antiinflammatory; vasotropic; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antiulcer; antithyroid; anti allergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;
KW murine; gene; ss.
XX
XX Mus musculus.
XX
XX WO2004040000-A2.
XX
XX 13-MAY-2004.
XX
XX 09-SEP-2003; 2003WO-US028226.
XX
XX 09-SEP-2002; 2002US-0409303P.
PR 09-APR-2003; 2003US-0461329P.
XX
XX (PRIM-) PRIMAL INC.
XX
XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vaesiliatis D, Zeng H;
PI
XX WPI; 2004-390329/36.
DR P-PSDB; ADO29633.
XX
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 1428; 542pp; English.
PS
XX
XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC

CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1068 BP; 205 A; 319 C; 267 G; 277 T; 0 U; 0 Other;
Query Match 50.0%; Score 21; DB 12; Length 1068;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAG 37
||||| | | | | | | | | | | | | | | | | | | | |
Db 1015 GCCACGCCCCGCCGACGTCGTACAAATGACTACGAAG 1051
XX
RESULT 84
ADL62421
ID ADL62421 standard; DNA; 1096 BP.
XX
AC ADL62421;
XX
XX 20-MAY-2004 (first entry)
XX Human ovarian cancer DNA marker #20633.
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX Homo sapiens.
OS
XX WO200170979-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US009126.
PF
XX 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lee J, Lillie J;
PI
XX WPI; 2001-611502/70.
DR
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
XX Disclosure; SEQ ID NO 20633; 106pp; English.
PS
XX The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a

CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1096 BP; 367 A; 215 C; 286 G; 224 T; 0 U; 4 Other;

Query Match 50.0%; Score 21; DB 5; Length 1096;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 5 AGCTTGGCCGAGCCCTTCAAAAGACTATGAAGTAAA 41
DB 437 AAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAAA 473

RESULT 85
ADL62598
ID ADL62598 standard; DNA; 1386 BP.
XX AC ADL62598;
XX DT 20-MAY-2004 (first entry)
XX DE Human ovarian cancer DNA marker #20810.
XX KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX OS Homo sapiens.
XX PN WO200170979-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009126.
XX PR 21-MAR-2000; 2000US-0191031P.
XX PR 25-MAY-2000; 2000US-0207124P.
XX PR 15-JUN-2000; 2000US-0211940P.
XX PR 07-JUL-2000; 2000US-0216820P.
XX PR 25-JUL-2000; 2000US-0220661P.
XX PR 21-DEC-2000; 2000US-0257672P.
XX PA (WILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Lee J, Lillie J;
XX DR WPI; 2001-611502/70.
XX PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
XX PT cancer cells as compared to their normal non-cancerous ovarian cells are
XX PT used to characterize stage, grade, histological type of ovarian cancer.
XX PS Disclosure; SEQ ID NO 20810; 106pp; English.
XX CC The invention relates to nucleic acid markers which are overexpressed in
XX CC ovarian cancer cells as compared to their expression in normal (i.e. non-
XX CC cancerous) ovarian cells. The invention also relates to polypeptides

CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
XX Sequence 1386 BP; 457 A; 287 C; 338 G; 300 T; 0 U; 4 Other;

Query Match 50.0%; Score 21; DB 5; Length 1386;
Best Local Similarity 73.0%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 5 AGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
DB 566 AAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAAA 602

RESULT 86
ADS10225
ID ADS10225 standard; DNA; 1566 BP.
XX AC ADS10225;
XX DT 16-DEC-2004 (first entry)
XX DE Human therapeutic DNA - SEQ ID 462.
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnary;
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX OS Homo sapiens.
XX PN WO2004080148-A2.
XX PD 23-SEP-2004.
XX PF 30-SEP-2003; 2003WO-US030720.
XX PR 02-OCT-2002; 2002US-0416186P.
XX PA (NUVE-) NUVELO INC.
XX PI Tang YT, Asundi V, Ren F, Zhang J, Zhang J, Wehrman T, Wang Z, Ma Y;
XX PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX WPI; 2004-668857/65.
XX DR P-PSDB; ADS10909.
XX PT New polynucleotide, useful in preparing a composition for diagnosing or


```
XX
SQ Sequence 3870 BP; 1071 A; 924 C; 938 G; 937 T; 0 U; 0 Other;
  Query Match          50.0%; Score 21; DB 13; Length 3870;
  Best Local Similarity 73.0%; Pred. No. 1.7e+02;
  Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAACTATAA 41
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 AAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAAA 646

RESULT 91
ADI22610/c
ID ADI22610 standard; DNA; 4605 BP.
XX
AC ADI22610;
XX
DT 22-APR-2004 (first entry)
XX
DE Human liver differentially expressed cDNA seq id 420.
XX
KW hepatotropic; antiinflammatory; virucide; cytostatic; gene therapy;
KW differentially gene expression; liver; toxin; liver disorder;
KW biliary cirrhosis; X-linked adrenoleukodystrophy; Zellweger syndrome;
KW hepatorenal syndrome; hepatitis; hepatocarcinoma; metabolic response;
KW toxicological response; ss; EST; expressed sequence tag; human.
XX
OS Homo sapiens.
XX
PN US2003165854-A1.
XX
PD 04-SEP-2003.
XX
PF 05-DEC-2001; 2001US-00006285.
XX
PR 05-DEC-2000; 2000US-0251986P.
XX
PA (CUNN/) CUNNINGHAM M J.
PA (KASE/) KASER M R.
XX
PI Cunningham MJ, Kaser MR;
XX
DR WPI; 2003-863697/80.
XX
PT New combination comprising a number of cDNAs that are differentially
PT expressed in a liver treated with a toxin, useful for diagnosing, staging
PT or treating liver disorders (e.g. biliary cirrhosis, hepatitis or
PT hepatocarcinoma).
XX
PS Claim 1; SEQ ID NO 420; 28pp; English.
XX
CC The invention describes a combination comprising a number of cDNAs that
CC are differentially expressed in a liver treated with a toxin and are
CC selected from any of the 514 cDNAs listed in the specification, or their
CC complements. The combination is useful in diagnosing, staging or treating
CC a liver disorder (e.g. biliary cirrhosis, X-linked adrenoleukodystrophy,
CC Zellweger syndrome, hepatorenal syndrome, hepatitis or hepatocarcinoma),
CC in monitoring diagnostic and therapeutic applications, in detecting
CC metabolic and toxicological responses, and in elucidating drug mechanism
CC of action. This sequence represents a cDNA differentially expressed in
CC liver tissues in response to treatment with a toxin.
XX
SQ Sequence 4605 BP; 1065 A; 1206 C; 1133 G; 1201 T; 0 U; 0 Other;
  Query Match          50.0%; Score 21; DB 10; Length 4605;
  Best Local Similarity 73.0%; Pred. No. 1.8e+02;
  Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAACTATAA 41
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3959 AAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAAA 3923

us-10-600-816-30.rng
RESULT 92
ADQ22756
ID ADQ22756 standard; DNA; 6346 BP.
XX
AC ADQ22756;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5576.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 5576; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 6346 BP; 1533 A; 1636 C; 1775 G; 1351 T; 0 U; 51 Other;
  Query Match          50.0%; Score 21; DB 12; Length 6346;
  Best Local Similarity 73.0%; Pred. No. 1.9e+02;
  Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAACTATAA 41
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 659 AAGTTTCACAGAGTCCTTCAAAAGACTCTGAAGAAAA 695

RESULT 93
ADC37602
ID ADC37602 standard; DNA; 7174 BP.
XX
AC ADC37602;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human nucleic acid associated protein NAAP-9 coding sequence.
XX
KW Human; nucleic acid associated protein; NAAP; cytostatic;
```

```
KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;
KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory;
KW thymetmetic; gene therapy; cell proliferative disorder; cancer;
KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;
KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
KW gene; ds.
XX
XX Homo sapiens.
XX
XX WO2003046151-A2.
XX
XX 05-JUN-2003.
XX
XX 26-NOV-2002; 2002WO-US038445.
XX
XX 27-NOV-2001; 2001US-0333925P.
XX
XX 07-DEC-2001; 2001US-0340477P.
XX
XX 14-DEC-2001; 2001US-0340362P.
XX
XX 18-DEC-2001; 2001US-0342002P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Baughn MR, Becha SD, Bhatia U, Blake JJ, Burford N, Burrill JD,
XX Chang H, Chawla NK, Elliott VS, Emerling BM, Forsythe IJ, Gandhi AR,
XX Gietzen KJ, Gorvad AE, Griffin JA, Hafalia AJA, Jackson JL, Ho A,
XX Ison CH, Jackson AA, Jiang X, Jin P, Kable AE, Khare R, Lal PG,
XX Lee EA, Lee S, Lee SY, Li JX, Lu DAM, Ramkumar J, Richardson TW,
XX Sprague WM, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H,
XX Zheng W;
XX WPI; 2003-513642/48.
XX
XX P-PSDB; ADC37542.
XX
XX New human nucleic acid associated proteins (NAAP), useful for diagnosing,
XX treating and preventing diseases or conditions associated with the
XX aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
XX infections.
XX
XX Claim 5; SEQ ID NO 69; 383pp; English.
XX
XX The present invention relates to novel human nucleic acid associated
XX protein (NAAP) (I; ADC37534-ADC37593) and their coding sequences ({}). The
XX NAAPs and their coding sequences are useful in diagnosing, treating and
XX preventing diseases or conditions associated with the decreased
XX expression or over expression of NAAP, such as cell proliferative (e.g.
XX cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
XX disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and
XX developmental (e.g. Hypothyroidism, Cushing's syndrome) disorders, or
XX infections. These are also useful in assessing the effects of exogenous
XX compounds on the expression of nucleic acid and amino acid sequences of
XX NAAP.
XX
XX Sequence 7174 BP; 1812 A; 1991 C; 1933 G; 1438 T; 0 U; 0 Other;
XX
XX Query Match 50.0%; Score 21; DB 10; Length 7174;
XX Best Local Similarity 73.0%; Pred. No. 2e+02;
XX Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 6 CCCTTGGCGGAGCCCTTACAAAGACTATGAGTAAAG 42
XX 460 CTCGTGGCGGAGGATCACAATTACAAATGAGTAAAG 496
XX
XX Db
XX
XX RESULT 94
XX ADA02708/c
XX ID ADA02708 standard; DNA; 96589 BP.
XX
XX AC ADA02708;
XX
XX DT 06-NOV-2003 (first entry)
XX
XX KW Human ZFXH1B carcinoma associated gene, SEQ ID NO:1226.
XX
XX DE
```

```
XX
XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO2003057146-A2.
XX
XX 17-JUL-2003.
XX
XX 26-DEC-2002; 2002WO-US041414.
XX
XX 26-DEC-2001; 2001US-00035832.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-587068/55.
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX
XX Claim 1; SEQ ID NO 1226; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX oncogenic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukaemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed human CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 96589 BP; 29004 A; 18509 C; 19494 G; 29562 T; 0 U; 20 Other;
XX
XX Query Match 50.0%; Score 21; DB 9; Length 96589;
XX Best Local Similarity 73.0%; Pred. No. 3.5e+02;
XX Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 5 AGCTTGGCGGAGCCCTTACAAAGACTATGAGTAAAG 41
XX 45448 ACTCTCTGCCGCGCTGTACAGAAATTATGAGAAAA 45412
XX
XX Db
XX
XX RESULT 95
XX ADB72446/c
XX ID ADB72446 standard; DNA; 96589 BP.
XX
XX AC ADB72446;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE Human ZFXH1B gene.
XX
XX KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
XX cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
XX KW
```


Query Match 49.0%; Score 20.6; DB 9; Length 457;
Best Local Similarity 85.2%; Pred. No. 1.6e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
|||||
Db 101 GAGCCCTTACAAAGACTATGAAGCACCANA 75

RESULT 100
ADL45933
ID ADL45933 standard; DNA; 682 BP.
XX
AC ADL45933;
XX
DT 20-MAY-2004 (first entry)
XX
XX Human ovarian cancer DNA marker #19823.

XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
XX Homo sapiens.
XX WO200170979-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009126.
XX
XX 21-MAR-2000; 2000US-0191031P.
XX
XX 25-MAY-2000; 2000US-0207124P.
XX
XX 15-JUN-2000; 2000US-0211940P.
XX
XX 07-JUL-2000; 2000US-0216820P.
XX
XX 25-JUL-2000; 2000US-0220661P.
XX
XX 21-DEC-2000; 2000US-0257672P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
XX Lee J, Lillie J;
XX
XX WPI; 2001-611502/70.
XX
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
XX cancer cells as compared to their normal non-cancerous ovarian cells are
XX used to characterize stage, grade, histological type of ovarian cancer.
XX
XX Disclosure; SEQ ID NO 19823; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in
XX ovarian cancer cells as compared to their expression in normal (i.e. non-
XX cancerous) ovarian cells. The invention also relates to polypeptides
XX encoded by the markers, antibodies that selectively bind to the
XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk
XX of developing ovarian cancer involving inhibiting expression of a gene
XX corresponding to a marker of the invention and a method of treating a
XX patient afflicted with ovarian cancer comprising providing to cells of
XX the patient an antisense oligonucleotide complementary to a marker of the
XX invention. The markers are useful for assessing if a patient is afflicted
XX with ovarian cancer, which involves comparing the level of expression of
XX a marker in a patient sample and a normal level of expression of the
XX marker in a control non-ovarian cancer sample. A difference between the
XX expression levels indicates ovarian cancer. The level of expression of a
XX marker corresponds to a secreted protein or to a transcribed
XX polynucleotide or its portion. The level of expression of the marker is
XX assessed by detecting the presence in the sample; a protein or protein
XX fragment corresponding to the marker. The presence of protein or protein
XX fragment is detected using an antibody that specifically binds with the
XX protein or protein fragment. Alternatively, the level of expression of
XX the marker is assessed by detecting the presence of a transcribed
XX polynucleotide which anneals with the marker or anneals with a portion of
XX the polynucleotide comprising the marker, under stringent conditions. The
XX marker is also used for monitoring the progression of ovarian cancer in a

CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 682 BP; 235 A; 112 C; 115 G; 209 T; 0 U; 11 Other;

Query Match 49.0%; Score 20.6; DB 5; Length 682;
Best Local Similarity 85.2%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
|||||
Db 436 GAGCCCTTACAAAGACTATGAAGCACCANA 462

Search completed: June 10, 2006, 15:32:20
Job time : 351.45 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:19:28 ; Search time 5182.8 Seconds
(without alignments)
453.155 Million cell updates/sec

Title: US-10-600-816-30
Perfect score: 42
Sequence: 1 gccacgctggccgagcc.....acaaagactgaagtaag 42

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : EST:
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	158	2	B063311 IL3-UT011
C 2	42	100.0	307	7	B818833 CM2-BN030
C 3	42	100.0	341	7	B818817 CM2-BN030
C 4	42	100.0	384	7	B818835 CM2-BN030
C 5	42	100.0	396	7	B818828 CM2-BN030
C 6	42	100.0	404	7	B818840 CM2-BN030
C 7	42	100.0	424	7	B818805 CM2-BN030
C 8	42	100.0	453	5	CD672930 fgl18a05.Y
C 9	42	100.0	458	7	B818871 CM2-BN030
C 10	42	100.0	466	7	B818859 CM2-BN030
C 11	42	100.0	467	7	B818852 CM2-BN030
C 12	42	100.0	571	3	BP253378 BP253378
C 13	42	100.0	577	4	CB152636 K-EST0209
C 14	42	100.0	580	3	BP265234 BP265234
C 15	42	100.0	580	3	BP268151 BP268151
C 16	42	100.0	582	3	BP256144 BP256144
C 17	42	100.0	582	3	BP363758 BP363758
C 18	42	100.0	584	3	BP256147 BP256147
C 19	42	100.0	621	4	CB130950 K-EST0180

C 531	20.6	49.0	898	3	BQ646878	604	20.4	48.6	539	3	BM839320	BM839320 K-EST0116
C 532	20.6	49.0	899	3	BUI189396	605	20.4	48.6	539	9	DA796970	DA796970 DA796970
C 533	20.6	49.0	900	3	DR6664118	606	20.4	48.6	539	9	DA989776	DA989776 DA989776
C 534	20.6	49.0	909	3	BUS08936	607	20.4	48.6	540	9	DA297709	DA297709 DA297709
C 535	20.6	49.0	915	10	DR672682	608	20.4	48.6	540	9	DA773480	DA773480 DA773480
C 536	20.6	49.0	919	7	BE885732	609	20.4	48.6	540	9	DA867860	DA867860 DA867860
C 537	20.6	49.0	922	10	DR671573	610	20.4	48.6	541	9	DA438005	DA438005 DA438005
C 538	20.6	49.0	925	3	BUI65566	611	20.4	48.6	542	9	DS001254	DS001254 DS001254
C 539	20.6	49.0	930	10	DR630257	612	20.4	48.6	542	9	DS122090	DS122090 DS122090
C 540	20.6	49.0	939	10	DR665600	613	20.4	48.6	543	9	DA575647	DA575647 DA575647
C 541	20.6	49.0	941	4	BX380892	614	20.4	48.6	543	9	DS258069	DS258069 DS258069
C 542	20.6	49.0	948	3	BM802535	615	20.4	48.6	544	9	DA276855	DA276855 DA276855
C 543	20.6	49.0	954	1	AL550675	616	20.4	48.6	544	9	DA651258	DA651258 DA651258
C 544	20.6	49.0	970	12	CC426405	617	20.4	48.6	546	9	DA758232	DA758232 DA758232
C 545	20.6	49.0	978	10	DR753481	618	20.4	48.6	548	9	DA353888	DA353888 DA353888
C 546	20.6	49.0	989	3	BM803520	619	20.4	48.6	548	9	DA715772	DA715772 DA715772
C 547	20.6	49.0	1027	3	BM903886	620	20.4	48.6	549	9	DA120549	DA120549 DA120549
C 548	20.6	49.0	1048	13	CL988825	621	20.4	48.6	549	9	DB129337	DB129337 DB129337
C 549	20.6	49.0	1079	1	AL578222	622	20.4	48.6	550	10	DT673629	DT673629 DT673629
C 550	20.6	49.0	1097	2	BG290255	623	20.4	48.6	551	9	DA932100	DA932100 DA932100
C 551	20.6	49.0	1139	6	C6060981	624	20.4	48.6	554	9	DA274458	DA274458 DA274458
C 552	20.6	49.0	1689	14	DQ053049	625	20.4	48.6	554	9	DB010322	DB010322 DB010322
C 553	20.6	49.0	2048	6	CR602997	626	20.4	48.6	555	9	DA180987	DA180987 DA180987
C 554	20.6	49.0	2752	6	BC014615	627	20.4	48.6	556	9	DA495232	DA495232 DA495232
C 555	20.6	49.0	2785	6	CR857698	628	20.4	48.6	556	9	DB104085	DB104085 DB104085
C 556	20.6	49.0	5305	6	BC042194	629	20.4	48.6	557	9	DA159711	DA159711 DA159711
C 557	20.4	48.6	196	13	CL655601	630	20.4	48.6	557	9	DA295987	DA295987 DA295987
C 558	20.4	48.6	239	7	B5884230	631	20.4	48.6	557	9	DA355653	DA355653 DA355653
C 559	20.4	48.6	253	10	DT911113	632	20.4	48.6	557	9	DA749462	DA749462 DA749462
C 560	20.4	48.6	255	1	AJ769677	633	20.4	48.6	557	9	DB141283	DB141283 DB141283
C 561	20.4	48.6	319	9	DA901604	634	20.4	48.6	558	9	DA724630	DA724630 DA724630
C 562	20.4	48.6	342	4	C81714	635	20.4	48.6	558	9	DA739090	DA739090 DA739090
C 563	20.4	48.6	371	5	CD700543	636	20.4	48.6	559	9	DA411398	DA411398 DA411398
C 564	20.4	48.6	397	9	DA944834	637	20.4	48.6	559	9	DA589646	DA589646 DA589646
C 565	20.4	48.6	404	10	DT053430	638	20.4	48.6	559	11	AZ087885	AZ087885 AZ087885
C 566	20.4	48.6	429	9	DN619561	639	20.4	48.6	561	9	DA824486	DA824486 DA824486
C 567	20.4	48.6	453	4	BY290880	640	20.4	48.6	561	9	DA875830	DA875830 DA875830
C 568	20.4	48.6	458	10	DR648792	641	20.4	48.6	562	9	DA052390	DA052390 DA052390
C 569	20.4	48.6	459	10	H01408	642	20.4	48.6	562	9	DA070759	DA070759 DA070759
C 570	20.4	48.6	460	9	DA123254	643	20.4	48.6	562	9	DA410342	DA410342 DA410342
C 571	20.4	48.6	460	9	DA184914	644	20.4	48.6	562	9	DA451541	DA451541 DA451541
C 572	20.4	48.6	467	9	DB015241	645	20.4	48.6	562	13	CW084394	CW084394 104 427 1
C 573	20.4	48.6	472	10	DR623909	646	20.4	48.6	563	9	DA451844	DA451844 DA451844
C 574	20.4	48.6	475	13	CZ362823	647	20.4	48.6	563	9	DA750855	DA750855 DA750855
C 575	20.4	48.6	484	2	BM153678	648	20.4	48.6	564	9	DA450701	DA450701 DA450701
C 576	20.4	48.6	486	9	DA197929	649	20.4	48.6	564	9	DA782343	DA782343 DA782343
C 577	20.4	48.6	490	9	DA740499	650	20.4	48.6	565	9	DA199685	DA199685 DA199685
C 578	20.4	48.6	490	9	DA741430	651	20.4	48.6	566	9	DA452973	DA452973 DA452973
C 579	20.4	48.6	499	13	DU101816	652	20.4	48.6	567	9	DA116268	DA116268 DA116268
C 580	20.4	48.6	500	9	DA741390	653	20.4	48.6	567	9	DA208058	DA208058 DA208058
C 581	20.4	48.6	500	9	DA741357	654	20.4	48.6	567	10	DR609565	DR609565 EST999693
C 582	20.4	48.6	503	9	DA889098	655	20.4	48.6	567	10	DA844736	DA844736 DA844736
C 583	20.4	48.6	504	1	AJ497845	656	20.4	48.6	568	9	DA021703	DA021703 DA021703
C 584	20.4	48.6	505	8	CN365809	657	20.4	48.6	568	9	DA397469	DA397469 DA397469
C 585	20.4	48.6	505	9	DA030327	658	20.4	48.6	569	9	DA577068	DA577068 DA577068
C 586	20.4	48.6	506	9	DA554409	659	20.4	48.6	569	9	DA922562	DA922562 DA922562
C 587	20.4	48.6	513	9	DA715433	660	20.4	48.6	569	9	DA386805	DA386805 DA386805
C 588	20.4	48.6	516	10	DR623216	661	20.4	48.6	570	9	DB051104	DB051104 DB051104
C 589	20.4	48.6	521	9	DA785287	662	20.4	48.6	570	9	DB098316	DB098316 DB098316
C 590	20.4	48.6	522	9	DA739803	663	20.4	48.6	570	9	DA297667	DA297667 DA297667
C 591	20.4	48.6	525	9	DA553880	664	20.4	48.6	571	9	DB164397	DB164397 DA164397
C 592	20.4	48.6	525	9	DA578817	665	20.4	48.6	571	9	BI427518	BI427518 sah81b07
C 593	20.4	48.6	525	9	DA850979	666	20.4	48.6	573	2	DA391956	DA391956 DA391956
C 594	20.4	48.6	525	9	DA923933	667	20.4	48.6	573	9	DA928209	DA928209 DA928209
C 595	20.4	48.6	526	1	AL598146	668	20.4	48.6	573	9	DA151492	DA151492 DA151492
C 596	20.4	48.6	526	9	DA205495	669	20.4	48.6	574	9	DA274851	DA274851 DA274851
C 597	20.4	48.6	527	10	DR625202	670	20.4	48.6	574	9	DA564521	DA564521 DA564521
C 598	20.4	48.6	531	9	DA151114	671	20.4	48.6	574	9	DA690472	DA690472 DA690472
C 599	20.4	48.6	532	9	DA745231	672	20.4	48.6	574	9	DB146555	DB146555 DB146555
C 600	20.4	48.6	533	9	DA418152	673	20.4	48.6	574	9	DB160985	DB160985 DB160985
C 601	20.4	48.6	536	8	CN365808	674	20.4	48.6	574	9		
C 602	20.4	48.6	536	9	DA621797	675	20.4	48.6	574	9		
C 603	20.4	48.6	537	9	DA391932	676	20.4	48.6	574	9		

677	20.4	48.6	575	9	DA752675	DA752675	C 750	20.4	48.6	661	8	CX053448
678	20.4	48.6	575	9	DB016042	DB016042	C 751	20.4	48.6	662	9	DN279867
679	20.4	48.6	575	9	DB198734	DB198734	752	20.4	48.6	662	13	CW176274
680	20.4	48.6	575	5	CD886611	G118.102M	753	20.4	48.6	663	8	CN365807
681	20.4	48.6	576	9	DA286467	DA286467	C 754	20.4	48.6	667	13	CZ648685
682	20.4	48.6	576	9	DA387063	DA387063	755	20.4	48.6	669	8	CV719946
683	20.4	48.6	576	9	DA924105	DA924105	756	20.4	48.6	672	8	CX079095
684	20.4	48.6	577	9	DB152847	DB152847	757	20.4	48.6	673	10	DR632354
685	20.4	48.6	579	9	DA296626	DA296626	758	20.4	48.6	674	10	DR613673
686	20.4	48.6	580	9	DA204059	DA204059	759	20.4	48.6	676	2	B1768647
687	20.4	48.6	580	9	DA389723	DA389723	760	20.4	48.6	679	5	CF125451
688	20.4	48.6	580	9	DA389948	DA389948	761	20.4	48.6	679	8	CX075282
689	20.4	48.6	580	10	DR622902	EST101303	762	20.4	48.6	680	10	DR628789
690	20.4	48.6	580	10	DR632404	EST102253	763	20.4	48.6	681	5	CF125357
691	20.4	48.6	581	9	DA151584	DA151584	764	20.4	48.6	681	10	DR624097
692	20.4	48.6	581	9	DA436817	DA436817	765	20.4	48.6	682	10	DR623515
693	20.4	48.6	581	9	DB259366	DB259366	766	20.4	48.6	682	10	DR623515
694	20.4	48.6	582	9	DA301416	DA301416	C 767	20.4	48.6	683	8	CV871575
695	20.4	48.6	584	9	DA863767	DA863767	768	20.4	48.6	684	10	DR625159
696	20.4	48.6	585	9	DA292745	DA292745	769	20.4	48.6	686	14	AG121405
697	20.4	48.6	585	9	DA345706	DA345706	770	20.4	48.6	686	13	CW309218
698	20.4	48.6	586	2	BG706909	602672082	771	20.4	48.6	698	14	EX209168
699	20.4	48.6	586	9	CX554931	Yds93f12	772	20.4	48.6	700	10	DR624004
700	20.4	48.6	586	9	DA988895	DA988895	773	20.4	48.6	701	7	BE731836
701	20.4	48.6	587	9	DA116506	DA116506	C 774	20.4	48.6	710	12	CC717906
702	20.4	48.6	587	9	DA128497	DA128497	775	20.4	48.6	711	10	DR630145
703	20.4	48.6	587	9	DA281071	DA281071	776	20.4	48.6	711	10	BZ038465
704	20.4	48.6	589	9	DA497274	DA497274	C 777	20.4	48.6	711	13	CW309219
705	20.4	48.6	591	9	DA771098	DA771098	778	20.4	48.6	712	10	DR624004
706	20.4	48.6	591	9	DB163011	DB163011	779	20.4	48.6	712	12	CL193312
707	20.4	48.6	592	5	CF766301	CBS000248	780	20.4	48.6	715	9	DN280932
708	20.4	48.6	594	9	DA676094	DA676094	781	20.4	48.6	720	10	DR622588
709	20.4	48.6	594	9	DB031221	DB031221	782	20.4	48.6	724	10	DR624030
710	20.4	48.6	594	9	DB098497	DB098497	783	20.4	48.6	724	10	DR624158
711	20.4	48.6	595	9	DA204383	DA204383	784	20.4	48.6	728	1	AU130485
C 712	20.4	48.6	595	10	DR607747	EST937875	785	20.4	48.6	731	12	CL155428
713	20.4	48.6	599	9	DA408949	DA408949	786	20.4	48.6	735	4	CA766731
714	20.4	48.6	599	9	DA536446	DA536446	787	20.4	48.6	738	7	BE733723
C 715	20.4	48.6	599	11	AZ602905	1M0421H20	788	20.4	48.6	740	10	DR631677
716	20.4	48.6	600	9	DB164566	DB164566	789	20.4	48.6	745	10	DT846508
C 717	20.4	48.6	601	4	CA144169	CA144169	790	20.4	48.6	750	8	CN365811
718	20.4	48.6	604	9	DA111137	DA111137	C 791	20.4	48.6	751	3	BU444563
719	20.4	48.6	605	11	AZ736059	RFC1-24-1	792	20.4	48.6	758	13	CL486136
C 720	20.4	48.6	609	9	CX638278	UCRPT023	793	20.4	48.6	761	2	BG753122
721	20.4	48.6	609	10	DR624446	EST101457	794	20.4	48.6	766	10	DR629331
C 722	20.4	48.6	610	10	DR611911	DR611911	795	20.4	48.6	770	12	CC346619
723	20.4	48.6	613	9	DB031613	DB031613	796	20.4	48.6	775	5	CJ456626
724	20.4	48.6	615	9	DA530469	DA530469	797	20.4	48.6	776	10	DR622257
725	20.4	48.6	620	7	AW239216	xb37d12.Y	798	20.4	48.6	779	10	DV907757
726	20.4	48.6	626	8	CN365810	170004240	799	20.4	48.6	781	10	DR614730
727	20.4	48.6	627	2	BG722171	BG722171	C 800	20.4	48.6	784	4	CB570527
728	20.4	48.6	629	10	DR631628	DR631628	801	20.4	48.6	791	10	DR626277
729	20.4	48.6	630	9	DA373137	DA373137	802	20.4	48.6	806	2	B1548671
730	20.4	48.6	632	10	DR623304	EST101343	C 803	20.4	48.6	810	13	CL399468
731	20.4	48.6	632	10	DR630991	EST102111	804	20.4	48.6	817	8	CO403906
732	20.4	48.6	632	10	DR632360	EST102248	C 805	20.4	48.6	832	10	DT085020
733	20.4	48.6	633	8	CN365813	170004241	C 806	20.4	48.6	833	10	DT0850917
C 734	20.4	48.6	635	2	CV710305	UCRPT01.0	807	20.4	48.6	834	9	DN288236
735	20.4	48.6	642	2	BG715825	BG715825	C 808	20.4	48.6	835	13	DUI51339
736	20.4	48.6	642	10	DV840366	1801128.C	C 809	20.4	48.6	838	8	CV561813
737	20.4	48.6	643	4	BX498357	DRFZP779D	810	20.4	48.6	838	9	DA197292
C 738	20.4	48.6	650	8	CV719945	UCRCS08.0	811	20.4	48.6	839	10	DR628297
739	20.4	48.6	652	2	BG333305	BG333305	812	20.4	48.6	841	12	CC628320
740	20.4	48.6	652	8	CN365800	170004553	C 813	20.4	48.6	842	14	CT471732
741	20.4	48.6	653	5	CX937461	CX937461	814	20.4	48.6	849	10	DR626504
742	20.4	48.6	654	8	CV710306	UCRPT01.0	815	20.4	48.6	852	10	DR630381
743	20.4	48.6	657	8	CX053449	UCRCS09_6	816	20.4	48.6	854	3	BQ213153
744	20.4	48.6	657	13	CW081126	CW081126	817	20.4	48.6	858	3	BU526982
745	20.4	48.6	658	8	CN915192	030628ABO	818	20.4	48.6	859	9	DN511381
- 746	20.4	48.6	658	10	DR613425	EST100355	819	20.4	48.6	861	12	CC717912
C 747	20.4	48.6	659	4	CB168018	OTM602700	820	20.4	48.6	872	10	DV881394
748	20.4	48.6	660	8	CT000578	CT000578	821	20.4	48.6	886	12	BZ792501
749	20.4	48.6	661	4	CA858155	EST635410	822	20.4	48.6	889	3	BU528156

CX053448 UCRCS09.6
DN279867 1160928.M
CW176274 104.589.1
CN365807 170004240
CZ648685 OM_Ba019
CV719946 UCRCS08.0
CX079095 yds07e1T
DR632354 EST102248
DR613673 EST100380
B1768647 603056989
CF125451 UI-HF-EL0
CX075282 UCRCS08.4
DR628789 EST101891
CF125357 UI-HF-EL0
DR624097 EST101422
DR623515 EST101364
DV871575 LB02515.C
CX075281 UCRCS08.4
DR625159 EST101528
AG121405 Pan trogl
CW309218 104.798.1
EX209168 Danilo rer
DR624004 EST101413
BE731836 601565890
CC717906 OGLAI45TH
DR630145 EST102027
BZ038465 1J042811.
CW309219 104.798.1
DV876369 LB0255.CR
CL193312 104.416.1
DN280932 1162080.M
DR622588 EST101271
DR624030 EST101415
DR624158 EST101428
AU130485 AU130485
CL155428 104.341.1
CA766731 AF53-Rof-1
BE733723 601568240
DR631677 EST102180
DT846508 LB00473.C
CN365811 170005315
BU444563 603763630
CL486136 SATL.428
BG753122 602732394-
DR629331 EST101945
CC346619 OGNBT14TV
CJ456626 CJ456626
DR622257 EST101238
DV907757 LB0283.CR
DR614730 EST100485
CB570527 AGENCOURT
DR626277 EST101640
B1548671 603196841
CL399468 ZMMBB040
CO403906 AGENCOURT
DT085020 JGI_AIZM4
DV760917 FchrSQ01
DN288236 1242726.M
DUI51339 OG_Aba003
CV561813 UI-N-HLO-
DA197292 DA197292
DR628297 EST101842
CC628320 OGLAK44TV
CT471732 Susa scrofi
DR626504 EST101663
DR630381 EST102050
BQ213153 AGENCOURT
BU526982 AGENCOURT
DN511381 1242462.M
CC717912 OGLAI45TV
DV881394 LB02618.C
BZ792501 PUFQ79TD
BU528156 AGENCOURT


```

969 20.2 48.1 5433 14 AY406469
970 20.2 48.1 6003 6 AK029889
971 20.2 48.1 8207 6 AK147591
972 20 47.6 126 1 AA750976
973 20 47.6 162 14 CR254107
974 20 47.6 187 1 AA749619
975 20 47.6 235 2 BJ663789
976 20 47.6 266 7 BB049175
977 20 47.6 269 3 BU095930
978 20 47.6 274 8 CO362116
979 20 47.6 275 7 BB441961
980 20 47.6 277 7 BB197214
981 20 47.6 280 7 BB439695
982 20 47.6 301 7 BB125964
983 20 47.6 303 7 BB189370
984 20 47.6 311 2 BG931284
985 20 47.6 320 7 BB209177
986 20 47.6 321 7 BB460781
987 20 47.6 324 2 BG931869
988 20 47.6 326 7 BB321388
989 20 47.6 328 7 BB557115
990 20 47.6 331 2 BG930982
991 20 47.6 335 1 A1707155
992 20 47.6 349 7 AV754927
993 20 47.6 388 4 BY635699
994 20 47.6 400 9 DN802613
995 20 47.6 445 4 CA967782
996 20 47.6 457 5 CF140081
997 20 47.6 469 9 DN802398
998 20 47.6 480 10 DR455490
999 20 47.6 496 9 DR071068
1000 20 47.6 511 14 CT142141

ALIGNMENTS

RESULT 1
BI063311 158 bp mRNA linear EST 15-JUN-2001
LOCUS IL3-UT0117-300301-538-E01 UT0117 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI063311
ACCESSION BI063311.1 GI:14470838
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 158)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsuoka,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-UT0117-
300301-538-E01&t3=2001-03-30&t4=1)
Seq primer: puc 18 forward

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-UT0117-
300301-538-E01&t3=2001-03-30&t4=1)
Seq primer: puc 18 forward

FEATURES
source
AY406469 Mus muscu
AK029889 Mus muscu
AK147591 Mus muscu
AA750976 ISKP0130
CR254107 Forward s
AA749619 ISAJ0092
BJ663789 BJ663789
BB049175 BB049175
BU095930 tca-415 t
CO362116 RTK1.1 H0
BB441961 BB441961
BB197214 BB197214
BB439695 BB439695
BB125964 BB125964
BB189370 BB189370
BG931284 f33-558 f
BB209177 BB209177
BB460781 BB460781
BG931869 h122-158
BB321388 BB321388
BB557115 BB557115
BG930982 f33-1147
A1707155 T6796 MVA
AV754927 AV754927
BY635699 BY635699
DN802613 G.hir-15-
CA967782 CcLLO3a12
CF140081 UI-HF-CB0
DN802398 G.hir-15-
DR455490 CM034G09
DR071068 RTDK1.17_
CT142141 Sus scrof

High quality sequence stop: 158.
Location/Qualifiers
1. .158
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UT0117"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 100.0%; Score 42; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
DB 100 GCCACGCTTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 141
|||||

RESULT 2
BE818833/c
LOCUS CM2-BN0302-050700-256-c01 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818833
ACCESSION BE818833.1 GI:10251067
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 307)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsuoka,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-BN0302-050
700-256-c01&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 307.

FEATURES
source
1. .307
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products

```

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 42; DB 7; Length 307;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||
 Db 242 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 201

RESULT 3

BE818817/c
 LOCUS CM2-BN0302-050700-256-b07 BN0302 Homo sapiens cDNA, mRNA sequence. EST 21-SEP-2000
 DEFINITION
 ACCESSION BE818817
 VERSION BE818817.1 GI:10251051
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=CM2-BN0302-050700-256-b07&t3=2000-07-05&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 341.

FEATURES
 source
 1. 341
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0302"
 /note="Organ: breast, normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 100.0%; Score 42; DB 7; Length 341;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||
 Db 253 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 212

RESULT 4

BE818835/c
 LOCUS CM2-BN0302-050700-256-cl1 BN0302 Homo sapiens cDNA, mRNA sequence. EST 21-SEP-2000
 DEFINITION
 ACCESSION BE818835
 VERSION BE818835.1 GI:10251069
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=CM2-BN0302-050700-256-cl1&t3=2000-07-05&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 5
 High quality sequence stop: 384.

FEATURES
 source
 1. 384
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0302"
 /note="Organ: breast, normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Query Match 100.0%; Score 42; DB 7; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||
 Db 257 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 216

RESULT 5
 BE818828/c
 LOCUS CM2-BN0302-050700-256-a05 BN0302 Homo sapiens cDNA, mRNA sequence. EST 21-SEP-2000
 DEFINITION

FEATURES
 source
 1. 396
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0302"
 /note="Organ: breast, normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Query Match 100.0%; Score 42; DB 7; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||
 Db 257 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 216

RESULT 5
 BE818828/c
 LOCUS CM2-BN0302-050700-256-a05 BN0302 Homo sapiens cDNA, mRNA sequence. EST 21-SEP-2000
 DEFINITION

```

ACCESSION   BE818828
VERSION     BE818828.1  GI:10251062
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 396)
AUTHORS    Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED     10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-050
            700-256-f02&t3=2000-07-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 396.

FEATURES             source
     source
     1..396
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /dev_stage="Adult"
         /clone_lib="BN0302"
         /note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
         Site 2: SmaI; A mini-library was made by cloning products
         derived from ORESTES PCR (U.S. Letters Patent application
         No. 196,716 - Ludwig Institute for Cancer Research)
         profiles into the pUC 18 vector. Reverse transcription of
         tissue mRNA and cDNA amplification were performed under
         low stringency conditions."

ORIGIN
Query Match      100.0%; Score 42; DB 7; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 42
    ||||||||||||||||||||||||||||||||||||||||
Db  248  GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 207

RESULT 6
BE818840/c
LOCUS       CM2-BN0302-050700-256-f02 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818840
ACCESSION  BE818840.1  GI:10251074
VERSION    BE818840.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 404)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED     10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-050
            700-256-f02&t3=2000-07-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 396.

FEATURES             source
     source
     1..404
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /dev_stage="Adult"
         /clone_lib="BN0302"
         /note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
         Site 2: SmaI; A mini-library was made by cloning products
         derived from ORESTES PCR (U.S. Letters Patent application
         No. 196,716 - Ludwig Institute for Cancer Research)
         profiles into the pUC 18 vector. Reverse transcription of
         tissue mRNA and cDNA amplification were performed under
         low stringency conditions."

ORIGIN
Query Match      100.0%; Score 42; DB 7; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 42
    ||||||||||||||||||||||||||||||||||||||||
Db  244  GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 203

RESULT 7
BE818805/c
LOCUS       CM2-BN0302-040700-253-b10 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818805
ACCESSION  BE818805.1  GI:10251039
VERSION    BE818805.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 424)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED     10737800
COMMENT     Contact: Simpson A.J.G.

```

```

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-050
700-256-f02&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 404.

FEATURES             source
     source
     1..404
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /dev_stage="Adult"
         /clone_lib="BN0302"
         /note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
         Site 2: SmaI; A mini-library was made by cloning products
         derived from ORESTES PCR (U.S. Letters Patent application
         No. 196,716 - Ludwig Institute for Cancer Research)
         profiles into the pUC 18 vector. Reverse transcription of
         tissue mRNA and cDNA amplification were performed under
         low stringency conditions."

ORIGIN
Query Match      100.0%; Score 42; DB 7; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 42
    ||||||||||||||||||||||||||||||||||||||||
Db  244  GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 203

RESULT 7
BE818805/c
LOCUS       CM2-BN0302-040700-253-b10 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818805
ACCESSION  BE818805.1  GI:10251039
VERSION    BE818805.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 424)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongenseel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED     10737800
COMMENT     Contact: Simpson A.J.G.

```


Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-cm2-BN0302-040
700-253-b10&t3=2000-07-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 424.
Location/Qualifiers

FEATURES

1..424
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

/note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 42; DB 7; Length 424;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 250 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 209
|||||

RESULT 8

CD672930 453 bp mRNA linear EST 24-JUN-2003
LOCUS fg18a05.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
fg18a05 5', mRNA sequence.

ACCESSION CD672930 GI:32174661

VERSION EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 453)

Wistow, G., Bernstein, S.L., Ray, S., Wyatt, M.K., Behal, A.,
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.

Expressed sequence tag analysis of adult human iris for the NEIBank
Project: steroid-response factors and similarities with retinal

pigment epithelium

Mol. Vis. 8 (4), 185-195 (2002)

JOURNAL 12107412

PUBMED

COMMENT Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 18 row: a column: 05

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

1..453

/organism="Homo sapiens"

/mol_type="mRNA"

FEATURES

source

1..458

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"
/clone_lib="fg18a05"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Iris cDNA (Normalized): fg"

/note="Organ: Eye; Vector: pCMVSPORT6; A human iris
library (ix) was normalized by self-subtraction. One
portion of double stranded plasmid DNA representing the
library was linearized by NotI. This NotI digested library
was used as a template for biotinylated RNA synthesis
using SP6 RNA polymerase. Another portion of the double
stranded plasmid library was converted to single-stranded
circles in vitro using Gene II and Exonuclease III (Life
Technologies). Single-stranded DNA (1 mg) was hybridized
(Cot 500) with 41 mg of Bio-RNA and vector blocking
oligonucleotides. The hybridized Bio-RNA/ss-circles were
removed by streptavidin:phenol extraction. EST analysis
was performed on the library at the NIH Intramural
Sequencing Center (NISC)."

ORIGIN

Query Match 100.0%; Score 42; DB 5; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 56 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 97
|||||

RESULT 9

BE818871/c

LOCUS

DEFINITION CM2-BN0302-100700-259-f01 BN0302 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE818871

VERSION BE818871.1 GI:10251105

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 458)

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL 10737800

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-cm2-BN0302-100
700-259-f01&t3=2000-07-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 12

High quality sequence stop: 458.

Location/Qualifiers

1..458

/organism="Homo sapiens"

/mol_type="mRNA"

```

/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

ORIGIN

```

Query Match      100.0%; Score 42; DB 7; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGTAAAG 42
    |||||||
Db 264 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGTAAAG 223

```

RESULT 10

```

BE818859/c
LOCUS              466 bp      mRNA      linear      EST 21-SEP-2000
DEFINITION      CM2-BN0302-100700-259-c02 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BE818859
VERSION        BE818859.1 GI:10251093
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

```

REFERENCE

```

1 (bases 1 to 466)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

```

```

Shotgun sequencing of the human transcriptome with ORF expressed

```

```

sequence tags

```

```

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800

```

JOURNAL

PUBMED

```

COMMENT

```

```

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

```

```

Tel: +55-11-2704922

```

```

Fax: +55-11-2707001

```

```

Email: asimpson@ludwig.org.br

```

```

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-100
700-259-c02&t3=2000-07-10&t4=1)

```

```

Seq primer: puc 18 forward

```

```

High quality sequence start: 16

```

```

High quality sequence stop: 466.

```

FEATURES

source

```

1. 466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

```

```

/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

ORIGIN

```

Query Match      100.0%; Score 42; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGTAAAG 42
    |||||||
Db 267 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGTAAAG 226

```

RESULT 11

```

BE818852/c
LOCUS              467 bp      mRNA      linear      EST 21-SEP-2000
DEFINITION      CM2-BN0302-100700-259-a05 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BE818852
VERSION        BE818852.1 GI:10251086
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

```

REFERENCE

```

1 (bases 1 to 467)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

```

```

Shotgun sequencing of the human transcriptome with ORF expressed

```

```

sequence tags

```

```

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800

```

JOURNAL

PUBMED

```

COMMENT

```

```

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

```

```

Tel: +55-11-2704922

```

```

Fax: +55-11-2707001

```

```

Email: asimpson@ludwig.org.br

```

```

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-100
700-259-a05&t3=2000-07-10&t4=1)

```

```

Seq primer: puc 18 forward

```

```

High quality sequence start: 19

```

```

High quality sequence stop: 467.

```

FEATURES

source

```

1. 467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

```

```

/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

ORIGIN

```

Query Match      100.0%; Score 42; DB 7; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGTAAAG 42
    |||||||
Db 246 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAGTAAAG 205

```

RESULT 12
 BP253378
 LOCUS
 DEFINITION BP253378 Sugano cDNA library, kidney epithelial cell Homo sapiens
 EST. BP253378.1 GI:52135659
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 571)
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 PUBMED 15342556
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: ysuzuki@igc.jp.
 Location/Qualifiers
 1..571
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HRC03720"
 /tissue_type="kidney"
 /cell_type="epithelial cell"
 /clone_lib="Sugano cDNA library, kidney epithelial cell"

Query Match 100.0%; Score 42; DB 3; Length 571;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 ||||||||||||||||||||||||||||||||||||||||
 Db 522 GCCCAGCTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 563

RESULT 13
 CB152636
 LOCUS
 DEFINITION CB152636 577 bp mRNA linear EST 29-JAN-2003
 5', mRNA sequence.
 ACCESSION
 VERSION CB152636.1 GI:28137590
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 577)
 AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr
 Plate: 5 row: G column: 07
 High quality sequence stop: 577.
 Location/Qualifiers
 1..577
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L12JSHC0s1-5-G07"
 /sex="M"
 /cell_line="J-SHC"
 /lab_host="Top10F"
 /clone_lib="L12JSHC0s1"
 /note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including
 EcoRI site by treatment of T4 RNA ligase and the first
 strand cDNA was synthesized from oligo dT-selected mRNA by
 priming with dT-tailed vector. The dT-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 - 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by
 PCR reaction using vector region primer including T7
 promoter as 5' primer and N(dT)14 as 3' primer. The PCR
 products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized cDNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybrids by centrifuge, the subtraced cDNA
 libraries were constructed by transformation of the
 remaining DNA into competent cells E. coli Top10F with
 electroporation method."

Query Match 100.0%; Score 42; DB 4; Length 577;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 ||||||||||||||||||||||||||||||||||||||||
 Db 197 GCCCAGCTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 238

RESULT 14
 BP265234
 LOCUS
 DEFINITION BP265234 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
 JTH00322, mRNA sequence.
 ACCESSION
 VERSION BP265234.1 GI:52180465
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 580)
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 PUBMED 15342556

```

COMMENT      Contact: Yutaka Suzuki
              Department of Virology
              Institute of Medical Science, University of Tokyo
              4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
              Tel: 81-3-5449-5343
              Fax: 81-3-5449-5416
              Email: yuzuki@hgc.jp.

FEATURES     source
              1. .580
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="JTH00322"
                /tissue_type="thyroid"
                /cell_line="JTH"
                /clone_lib="Sugano cDNA library, thyroid JTH"
                /note="thyroid tumor"

ORIGIN
Query Match      100.0%; Score 42; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||||||||||||||||||||||||||||||||||
Db 327 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 368

RESULT 15
LOCUS      BP268151          580 bp      mRNA      linear      EST 16-SEP-2004
DEFINITION BP268151 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
              JTH09144, mRNA sequence.
ACCESSION  BP268151
VERSION     BP268151.1 GI:52183383
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 580)
            Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
            Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
            Genome Res. 14 (9), 1711-1718 (2004)
            15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Tel: 81-3-5449-5343
            Fax: 81-3-5449-5416
            Email: yuzuki@hgc.jp.
            Location/Qualifiers
              1. .582
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="HRC10671"
                /tissue_type="kidney"
                /cell_type="epithelial cell"
                /clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
Query Match      100.0%; Score 42; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||||||||||||||||||||||||||||||||||
Db 415 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 456

RESULT 17
LOCUS      BP363758          582 bp      mRNA      linear      EST 17-SEP-2004
DEFINITION BP363758 Sugano cDNA library, fetal lung fibroblast TIG Homo
              sapiens cDNA clone TIR02383, mRNA sequence.
ACCESSION  BP363758
VERSION     BP363758.1 GI:52293963
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 582)
            Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
            Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
            Genome Res. 14 (9), 1711-1718 (2004)
            15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo

```

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: ysuzuki@hgc.jp.

FEATURES

source
 1..582
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TIR02383"
 /tissue_type="lung"
 /cell_type="fibroblast"
 /cell_line="TIG"
 /dev_stage="fetal"
 /clone_lib="Sugano cDNA library, fetal lung fibroblast TIG"

ORIGIN

Query Match 100.0%; Score 42; DB 3; Length 582;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||
 Db 380 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 421

RESULT 18

BP256147
 LOCUS BP256147 Sugano cDNA library, kidney epithelial cell Homo sapiens
 DEFINITION BP256147 Sugano cDNA library, kidney epithelial cell Homo sapiens
 CNA clone HRC10676, mRNA sequence.
 VERSION BP256147
 EST. BP256147.1 GI:52171377

SOURCE

ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

1 (Bases 1 to 584)
 Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)

JOURNAL

PUBMED 15342556
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: ysuzuki@hgc.jp.

FEATURES

source
 1..584
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HRC10676"
 /tissue_type="kidney"
 /cell_type="epithelial cell"
 /clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN

Query Match 100.0%; Score 42; DB 3; Length 584;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||
 Db 416 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 457

RESULT 19

CB130950
 LOCUS CB130950 K-EST0180946 L12JSHC0 Homo sapiens cDNA clone L12JSHC0-5-D03 5',
 DEFINITION mRNA sequence.

ACCESSION

VERSION CB130950

KEYWORDS

SOURCE CB130950.1 GI:28095412

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 621)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.

TITLE

JOURNAL 21C Frontier Korean EST Project 2001

COMMENT

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 5 row: D column: 03

High quality sequence stop: 621.

FEATURES

source

1..621
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L12JSHC0-5-D03"
 /sex="M"
 /cell_line="J-SHC"
 /lab_hosts="Top10F"
 /clone_lib="L12JSHC0"
 /note="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including
 EcoRI site by treatment of T4 RNA ligase and the first
 strand cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F, by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Query Match 100.0%; Score 42; DB 4; Length 621;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||
 Db 314 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 355

RESULT 20

BE747107
 LOCUS BE747107 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3929706 5',
 DEFINITION mRNA sequence.

ACCESSION

VERSION BE747107.1 GI:10161099

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 630)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM762 row: 0 column: 19
High quality sequence stop: 628.
FEATURES source
1..630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-77-H12"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/notes="Organ: Ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

ORIGIN
Query Match 100.0%; Score 42; DB 7; Length 630;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 66 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 107
|||||

RESULT 21
BM843051
LOCUS K-EST0120649 S12SNU216 Homo sapiens cDNA clone S12SNU216-77-H12 5',
mRNA sequence.
ACCESSION BM843051
VERSION BM843051.1 GI:19199460
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 637)
AUTHORS Oh, J.H., Yang, J.O., Hahn, Y., Kim, M.R., Byun, S.S., Jeon, Y.J.,
Kim, J.M., Song, K.S., Noh, S.M., Kim, S., Yoo, H.S., Kim, Y.S. and
Kim, N.S.
TITLE Transcriptome analysis of human gastric cancer
JOURNAL Mamm. Genome 16 (12), 942-954 (2005)
PUBMED 16341674
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
```

```

FEATURES source
1..637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-77-H12"
/lab_hosts="Top10F"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/clone_lib="S12SNU216"
/notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

```

ORIGIN
Query Match 100.0%; Score 42; DB 3; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 59 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 100
|||||

RESULT 22
BF568108
LOCUS 602183908F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4300248 5',
mRNA sequence.
ACCESSION BF568108
VERSION BF568108.1 GI:11641526
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 699)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM158 row: 0 column: 01
High quality sequence stop: 638.
FEATURES source
1..699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-77-H12"
/lab_hosts="Top10F"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/clone_lib="S12SNU216"
/notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```


/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4300248"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_42"
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. "

ORIGIN

Query Match 100.0%; Score 42; DB 7; Length 699;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||
 Db 321 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 362
 |||||

RESULT 23

LOCUS BG251131 702 bp mRNA linear EST 13-FEB-2001
 DEFINITION 602364932F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473439 5',
 mRNA sequence.

ACCESSION BG251131 GI:12760947

VERSION BG251131

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

1 (bases 1 to 702)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Cloning by: Incyte Genomics, Inc.

Cloning distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM10295 row: 0 column: 08

High quality sequence stop: 688.

FEATURES

source

Location/Qualifiers

1..702

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4473439"

/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_90"

/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 42; DB 2; Length 702;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||
 Db 464 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 505
 |||||

RESULT 24

LOCUS BQ887352 851 bp mRNA linear EST 16-AUG-2002
 DEFINITION AGENCOURT 8678064 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381080
 5', mRNA sequence.

ACCESSION BQ887352

VERSION BQ887352.1 GI:22279366

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

1 (bases 1 to 851)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/BTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LUCM2570 row: 1 column: 09

High quality sequence stop: 579.

FEATURES

source

Location/Qualifiers

1..851

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6381080"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 42; DB 3; Length 851;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||

Db 182 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 223
 |||||

RESULT 25

LOCUS CA454746 865 bp mRNA linear EST 12-NOV-2002
 DEFINITION AGENCOURT 10763191 MAPcL Homo sapiens cDNA clone IMAGE:6721245 5',
 mRNA sequence.

ACCESSION CA454746

VERSION CA454746.1 GI:24904781

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS TITLE JOURNAL COMMENT

1 (bases 1 to 865)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Kristi A. Egland, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M14282 row: a column: 21

High quality sequence stop: 686.

FEATURES source

1..865
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721245"
/call_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCap"
/lab_host="EMDH10B"
/clone_lib="NAPCL"

/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 100.0%; Score 42; DB 4; Length 865;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 611 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 652

RESULT 26 BM011054

DEFINITION 603634744F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5421755 5',
mRNA sequence.

ACCESSION BM011054

VERSION BM011054.1 GI:16525408

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 877)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M1878 row: h column: 12
High quality sequence stop: 746.

FEATURES source

1..877
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5421755"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_47"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 42; DB 2; Length 877;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 184 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 225

RESULT 27

BQ958341

LOCUS

DEFINITION BQ958341 899 bp mRNA linear EST 21-AUG-2002
AGENCOURT 10037048 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6480481
5' RNA sequence.

ACCESSION BQ958341

VERSION BQ958341.1 GI:22373819

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 899)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M2660 row: f column: 02

High quality sequence stop: 636.

FEATURES source

1..899
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6480481"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 42; DB 3; Length 899;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 546 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 587
|||||

RESULT 28

BG831564
LOCUS 602765819F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4907885 5',
DEFINITION mRNA sequence.

ACCESSION BG831564

VERSION BG831564.1 GI:14179151

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 901)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM1813 row: e column: 06

High quality sequence stop: 663.

FEATURES

source

1. 901
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4907885"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."|

ORIGIN

Query Match 100.0%; Score 42; DB 2; Length 901;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||

Db 322 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 363
|||||

RESULT 29

BX372160/c

LOCUS

DEFINITION BX372160 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK011YF17 3-PRIME, mRNA sequence.

ACCESSION

VERSION BX372160

KEYWORDS

SOURCE EST.

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 918)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

COMMENT Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

6601.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0BAI040ZB07_CS03797_1&c=6601.f

FEATURES

source

Location/Qualifiers

1. 918

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="CS0DK011YF17"

/cell_type="HELA CELLS COT 25-NORMALIZED"

/clone_lib="HELA"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 42; DB 4; Length 918;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||

Db 335 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 294
|||||

RESULT 30

BUS43952

LOCUS

DEFINITION AGNCOURT 10333041 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6576152
5', mRNA sequence.

ACCESSION

VERSION BUS43952

KEYWORDS

SOURCE EST.

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 923)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2774 row: d column: 08
 High quality sequence stop: 641.
 Location/Qualifiers
 1. 923
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6576152"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC 40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 42; DB 3; Length 923;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 42
 |||||
 Db 546 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 587

RESULT 31

BX370558
 LOCUS BX370558 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DK011YF17 5-PRIME, mRNA sequence.
 ACCESSION BX370558
 VERSION BX370558.1 GI:30453927
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

1 (bases 1 to 962)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 6601.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAG0492C03_CS04627_1kc=6601.f

FEATURES

source
 Location/Qualifiers
 1. 962
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK011YF17"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"

/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 42; DB 4; Length 962;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 42
 |||||
 Db 248 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 289

RESULT 32

BX370558
 LOCUS BX370558 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DK011YF17 5-PRIME, mRNA sequence.
 ACCESSION BX370558
 VERSION BX370558.1 GI:19896479
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

1 (bases 1 to 1023)

NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-x@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2793 row: c column: 09

High quality sequence stop: 649.

Location/Qualifiers

FEATURES

source
 Location/Qualifiers
 1. 1023
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5754944"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC 115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 42; DB 3; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 42
 |||||
 Db 403 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAGCTAAAG 444

```

RESULT 33
CR597125      1830 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS
DEFINITION
full-length cDNA clone CS0DK011YF17 of HeLa cells Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR597125
VERSION
CR597125.1 GI:50477932
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1830)
AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
2 (bases 1 to 1830)
Genoscope.
Direct Submission
JOURNAL
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YF17"
/tissue type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match      100.0%; Score 42; DB 6; Length 1830;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1118 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1159
|||||

RESULT 34
CF780868      916 bp      mRNA      linear      EST 20-OCT-2003
LOCUS
DEFINITION
AGENCOURT_15939245 NIH_MGC_219 Homo sapiens cDNA clone
IMAGE:30523569 5', mRNA sequence.
ACCESSION
CF780868
VERSION
CF780868.1 GI:37740645
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 916)
AUTHORS
NIH-MGC http://mgs.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa

```

```

CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM601 row: 1 column: 10
High quality sequence stop: 651.
FEATURES
Location/Qualifiers
source
1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30523569"
/tissue type="Pooled Chondrosarcoma Tumor cells"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_219"
/note="Vector: pYX-Asc; Site1: EcoRI; Site2: NotI;
Library is oligo-dT primed and directionally cloned
Denatured RNA was size fractionated on a 1% agarose gel.
First strand cDNA synthesis was primed with oligo-dT
primer containing a NotI site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoRI adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. Average insert size
0.5-1kb. Adaptors 5'(AATTGCGCAGCG)3' and 5'd
(CCTCGGCGCG)3'. 3' linker sequence - GCGGCGCTGAGAGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTAACCCCTACATAAGGAG)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Average insert size 2-3kb.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match      97.6%; Score 41; DB 5; Length 916;
Best Local Similarity 97.6%; Pred. No. 3.9e-06;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 667 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 708
|||||

RESULT 35
BI193620      1125 bp      mRNA      linear      EST 10-JUL-2001
LOCUS
DEFINITION
602946519F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5089900 5',
mRNA sequence.
ACCESSION
BI193620
VERSION
BI193620.1 GI:14648640
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1125)
AUTHORS
NIH-MGC http://mgs.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1851 row: e column: 05
High quality sequence stop: 529.
FEATURES
Location/Qualifiers
source
1..1125

```

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IWAGS:5089900"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_42"
 /notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. |"

ORIGIN

Query Match 97.6%; Score 41; DB 2; Length 1125;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
 |||||||
 Db 294 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 334

RESULT 36
BE818831/c

LOCUS BE818831 365 bp mRNA linear EST 21-SEP-2000
 DEFINITION CM2-BN0302-050700-256-al2 BN0302 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE818831
 VERSION BE818831.1 GI:10251065
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

1 (bases 1 to 365)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

PUBMED

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-CM2-BN0302-050
 700-256-al2&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 14

High quality sequence stop: 365.

FEATURES

source

1..365
 Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BN0302"

/note="Organ: breast_normal; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 96.2%; Score 40.4; DB 7; Length 365;
 Best Local Similarity 97.6%; Pred. No. 5.9e-06;
 Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 42
 |||||||
 Db 267 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 226

RESULT 37

BE818807/c

LOCUS BE818807 449 bp mRNA linear EST 21-SEP-2000
 DEFINITION CM2-BN0302-040700-253-d02 BN0302 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE818807

VERSION BE818807.1 GI:10251041

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

1 (bases 1 to 449)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-CM2-BN0302-040
 700-253-d02&t3=2000-07-04&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 24

High quality sequence stop: 306.

FEATURES

source

1..449
 Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BN0302"

/note="Organ: breast_normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 96.2%; Score 40.4; DB 7; Length 449;

Best Local Similarity 97.6%; Pred. No. 6.1e-06;

Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 240 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 199
|||||

RESULT 38
BP267627 580 bp mRNA linear EST 16-SEP-2004
LOCUS BP267627 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
DEFINITION JTH07584, mRNA sequence.

ACCESSION BP267627
VERSION BP267627.1 GI:52182859
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556

COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yusuzuki@hgc.jp.

FEATURES
source
1. 580
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH07584"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/note="thyroid tumor"

ORIGIN
Query Match 96.2%; Score 40.4; DB 3; Length 580;
Best Local Similarity 97.6%; Pred. No. 6.4e-06;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 539 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 580
|||||

RESULT 39
BQ685729 940 bp mRNA linear EST 15-JUL-2002
LOCUS BQ685729 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251057
DEFINITION AGNCOURT_8344399 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251057
5', mRNA sequence.

ACCESSION BQ685729
VERSION BQ685729.1 GI:21811045
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1CM2395 row: b column: 18
High quality sequence stop: 538.
Location/Qualifiers
1. 940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6251057"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES
source
1. 940
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6251057"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 96.2%; Score 40.4; DB 3; Length 940;
Best Local Similarity 97.6%; Pred. No. 7e-06;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 687 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 728
|||||

RESULT 40
AUI40676 680 bp mRNA linear EST 05-AUG-2002
LOCUS AUI40676 PLACE4 Homo sapiens cDNA clone PLACE4000034 5', mRNA
DEFINITION AUI40676 PLACE4 Homo sapiens cDNA clone PLACE4000034 5', mRNA
sequence.

ACCESSION AUI40676
VERSION AUI40676.1 GI:11002197
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 680)
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Iehida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

TITLE Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes

JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560

COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

FEATURES

source
 Location/Qualifiers
 1. .680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE4000034"
 /tissue_type="placenta"
 /clone_lib="PLACE4"
 /note="Vector: pME18SFL3"

ORIGIN

Query Match 80.0%; Score 33.6; DB 1; Length 680;
 Best Local Similarity 87.8%; Pred. No. 0.0051;
 Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCCACCGCTTGGCGAGCCCTTACAAAGACTATGAGTAAA 41
 |||||
 Db 633 GCCACCGCTTGGCGAGCCCTTACAAAGACTATGAGTAA 673

RESULT 41

AA112374 367 bp mRNA linear EST 01-AUG-1997
 LOCUS z68e10.11 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
 DEFINITION IMAGE:563370 5', mRNA sequence.
 AA112374
 ACCESSION AA112374.1 GI:1664784
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
 TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 PUBMED 8889549

TITLE

Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 PUBMED 8889549

COMMENT

Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 1363 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 329.

FEATURES

source
 Location/Qualifiers
 1. .367
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:4593181"
 /db_xref="taxon:9606"
 /clone="IMAGE:563370"
 /sex="female"
 /dev_stage="HeLa S3 cell line"
 /lab_host="SOJR (kanamycin resistant)"
 /clone_lib="Stragene HeLa cell s3 937216"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dr. HeLa S3 epitheloid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR Vector. -5' adaptor sequence: 5' GAATTCGGCAG 3' -3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTT 3'."

ORIGIN

Query Match 73.8%; Score 31; DB 1; Length 367;
 Best Local Similarity 97.7%; Pred. No. 0.058;
 Matches 42; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 1 GCCACCGCTTGGCGAGCCCTTACAAAGACTATGAGTAAA 42
 |||||
 Db 116 GCCACCGCTTGGCGAGCCCTTACAAAGACTATGAGTAAA 158

RESULT 42

BE006062 583 bp mRNA linear EST 05-JUN-2000
 LOCUS RC2-BN0123-170300-011-c07 BN0123 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE006062
 ACCESSION BE006062.1 GI:8266295
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=RC2-BN0123-170

300-011-c07&t3=2000-03-17&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 24

High quality sequence stop: 583.

Location/Qualifiers

1. .583

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BN0123"

/note="Organ: breast normal; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

ORIGIN

Query Match 72.4%; Score 30.4; DB 7; Length 583;
 Best Local Similarity 96.9%; Pred. No. 0.11;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCCACCGCTTGGCGAGCCCTTACAAAGACTA 32
 |||||
 Db 552 GCCACCGCTTGGCGAGCCCTTACAAAGACTA 583

```

RESULT 43
DN875318
LOCUS      DN875318          606 bp      mRNA      linear      EST 21-APR-2005
DEFINITION cdna clone n838908.Y1 Dog eye cornea. Unnormalized (nad) Canis familiaris
ACCESSION  DN875318
VERSION     DN875318.1  GI:62845259
KEYWORDS   EST.
SOURCE     Canis familiaris (dog)
ORGANISM   Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE  1  (bases 1 to 606)
AUTHORS   Wistow,G.
TITLE     Neibank analysis of Dog cornea
JOURNAL   Unpublished (2005)
COMMENT   Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 38 row: g column: 08
            Seq primer: Universal M13 Reverse.
            Location/Qualifiers
                1..606
                /organism="Canis familiaris"
                /mol_type="mRNA"
                /strain="Beagle"
                /db_xref="taxon:9615"
                /clone="nad38908"
                /tissue_type="Cornea"
                /dev_stage="Adult"
                /lab_host="EMDH10B"
                /clone_lib="Dog eye cornea. Unnormalized (nad)"
                /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted
                from dog cornea tissue. A directionally cloned cDNA
                library in the pCMVSPORT6 vector (Invitrogen) was
                constructed at Bioserve Biotechnology (Laurel MD)
                essentially following the protocols of the SuperScript
                Plasmid System, full details of which are contained in the
                manufacturer's instruction manual
                (http://www.lifetech.com/). First strand synthesis was
                carried out using a Not I primer-adaptor
                [5'-pGACTAGTTCTAGATCGGCGGCCGCTT15-3']. cDNA was
                cloned in Not I/Sal I sites. EST analysis was performed at
                the NIH Intramural Sequencing Center (NISC). Analyzed data
                available through http://neibank.nei.nih.gov."

ORIGIN
Query Match      69.5%; Score 29.2; DB 9; Length 606;
Best Local Similarity 81.0%; Pred. No. 0.37;
Matches 34; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      286  GCCCAGCTCGGACTAGCCCTTACAAAGACTATGAAGGAGG 327

RESULT 44
BE548141
LOCUS      BE548141          912 bp      mRNA      linear      EST 09-AUG-2000
DEFINITION 601072396F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3458567 5',
            mRNA sequence.
ACCESSION  BE548141
VERSION     BE548141.1  GI:9776786
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1  (bases 1 to 912)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8449 row: p column: 24
            High quality sequence stop: 651.
            Location/Qualifiers
                1..912
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3458567"
                /tissue_type="cervical carcinoma cell line"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_12"
                /note="Organ: Cervix; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
                Average insert size 1.4 kb. Library prepared by Life
                Technologies."

ORIGIN
Query Match      67.6%; Score 28.4; DB 7; Length 912;
Best Local Similarity 82.1%; Pred. No. 0.88;
Matches 32; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy      4  CACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      620  CCCAGCGGCGAGCCCTTACAAAGACTATGAAGTCAAG 658

RESULT 45
AW390096
LOCUS      AW390096          528 bp      mRNA      linear      EST 04-FEB-2000
DEFINITION CM3-ST0179-051099-019-f07 ST0179 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW390096
VERSION     AW390096.1  GI:6894755
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1  (bases 1 to 528)
AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
TITLE     The FAPESP/LICR Human Cancer Genome Project
JOURNAL   Unpublished (1999)
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-ST0179-
            051099-019-f07&t3=1999-10-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 528.
            Location/Qualifiers

```

```

REFERENCE
AUTHORS   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
TITLE     1  (bases 1 to 912)
JOURNAL   NIH-MGC http://mgi.nci.nih.gov/.
COMMENT   National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8449 row: p column: 24
            High quality sequence stop: 651.
            Location/Qualifiers
                1..912
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3458567"
                /tissue_type="cervical carcinoma cell line"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_12"
                /note="Organ: Cervix; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
                Average insert size 1.4 kb. Library prepared by Life
                Technologies."

ORIGIN
Query Match      67.6%; Score 28.4; DB 7; Length 912;
Best Local Similarity 82.1%; Pred. No. 0.88;
Matches 32; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy      4  CACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      620  CCCAGCGGCGAGCCCTTACAAAGACTATGAAGTCAAG 658

RESULT 45
AW390096
LOCUS      AW390096          528 bp      mRNA      linear      EST 04-FEB-2000
DEFINITION CM3-ST0179-051099-019-f07 ST0179 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW390096
VERSION     AW390096.1  GI:6894755
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1  (bases 1 to 528)
AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
TITLE     The FAPESP/LICR Human Cancer Genome Project
JOURNAL   Unpublished (1999)
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-ST0179-
            051099-019-f07&t3=1999-10-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 528.
            Location/Qualifiers

```

```

source
1. 528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0179"
/notes="Organ: stomach; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 64.3%; Score 27; DB 7; Length 528;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCGAGCCCTTACAAA 27
| | | | | | | | | | | | | | | | | | | | |
Db 502 GCCACGCTTGCGCGAGCCCTTACAAA 528

RESULT 46
BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
LOCUS JTH08419, mRNA sequence.
DEFINITION BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
ACCESSION BP267918
VERSION BP267918.1 GI:52183150
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yezuki@hgc.jp.
Location/Qualifiers
1. 582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH08419"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"

```

```

ORIGIN
Query Match 64.3%; Score 27; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCGAGCCCTTACAAA 27
| | | | | | | | | | | | | | | | | | | | |
Db 556 GCCACGCTTGCGCGAGCCCTTACAAA 582

RESULT 47
BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
LOCUS JTH08419, mRNA sequence.
DEFINITION BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
ACCESSION BP267918
VERSION BP267918.1 GI:52183150
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yezuki@hgc.jp.
Location/Qualifiers
1. 582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH08419"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"

```

```

ORIGIN
Query Match 64.3%; Score 27; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCGAGCCCTTACAAA 27
| | | | | | | | | | | | | | | | | | | | |
Db 556 GCCACGCTTGCGCGAGCCCTTACAAA 582

RESULT 47
BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
LOCUS JTH08419, mRNA sequence.
DEFINITION BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
ACCESSION BP267918
VERSION BP267918.1 GI:52183150
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yezuki@hgc.jp.
Location/Qualifiers
1. 582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH08419"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"

```

```

BE815636/c
LOCUS BE815636 240 bp mRNA linear EST 21-SEP-2000
DEFINITION PM4-BN0172-010600-006-h11 BN0172 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE815636
VERSION BE815636.1 GI:10247870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 240)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-pm4-BN0172-010
600-006-h11&t3=2000-06-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 79.
Location/Qualifiers
1. 240
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0172"
/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

JOURNAL BE815636
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-pm4-BN0172-010
600-006-h11&t3=2000-06-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 79.
Location/Qualifiers
1. 240
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0172"
/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

FEATURES
source
1. 240
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0172"
/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

ORIGIN
Query Match 62.4%; Score 26.2; DB 7; Length 240;
Best Local Similarity 90.3%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 10 TGCCCGAGCCCTTACAAAGACTATGAAGTAA 40
| | | | | | | | | | | | | | | | | | | | |
Db 151 TTGCGGAGCCCTTACAAAGACTATGAATAA 121

RESULT 48
AA222951 336 bp mRNA linear EST 12-MAR-1997
LOCUS mw01h08.r1 GuayWoodford Beier mouse kidney day 7 Mus musculus cDNA
DEFINITION clone IMAGE:663327 5', mRNA sequence.
ACCESSION AA222951
VERSION AA222951.1 GI:1843213
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```

RESULT 56
DN020621
LOCUS
DEFINITION
  JGI CAAR2620.rev NIH XGC tropLiv1 Xenopus tropicalis cDNA clone
  IMAGE:7734970 3', mRNA sequence.
ACCESSION
DN020621
VERSION
KEYWORDS
SOURCE
  DN020621.1 GI:58864255
  EST.
  Xenopus tropicalis (western clawed frog)
  Xenopus tropicalis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
  Xenopodinae; Xenopus; Silurana.
  1 (bases 1 to 864)
  Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,
  Brokstein,P. and Lindquist,E.A.
  DOE Joint Genome Institute Xenopus tropicalis EST project
  Unpublished (2004)
  Other ESTs: JGI CAAR2620.fwd
  Contact: Lindquist,E.A., Richardson,P.
  DOE Joint Genome Institute
  2800 Mitchell Drive, Walnut Creek, CA 94598, USA
  Tel: 925 296 5600
  Fax: 925 296 5710
  Email: cdna@jgi-psf.org
  Tissue Procurement: Robert M. Grainger
  cDNA Library Preparation: Bruce Blumberg Laboratory, University of
  California, Irvine
  DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
  Clone Distribution: I.M.A.G.E. Consortium/LLNL;
  http://image.llnl.gov
  Naming Conventions: EST name is generated by the concatenation of
  the JGI Clone id and the direction of sequencing. The suffix '.rev'
  indicates a reverse sequencing read of the insert. It does not
  necessarily reflect the orientation of the insert.
  Poly-A: Based upon the presence of a run of 14 or more T residues
  at the beginning of the sequence, this clone was polyadenylated.
  The resulting Poly-T sequence has been removed.
  Plate: CAAR 0025 row: h column: 8
  High quality sequence stop: 751
  POLYA=Yes.
  Location/Qualifiers
    1..864
      /organism="Xenopus tropicalis"
      /mol_type="mRNA"
      /strain="N6 (Nigerian 6th generation inbred)"
      /db_xref="taxon:8164"
      /clone="IMAGE:7734970"
      /sex="male"
      /tissue_type="Liver"
      /dev_stage="Adult"
      /lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
      /clone_lib="NIH XGC tropLiv1"
      /note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
      library was prepared from 5 ug of poly A+ RNA by oligo-dT
      priming
      (5'-GAGAGAGAGAGAGAGACTAGTCTCGAGTCTTTTCTTTT-3')
      and Stratascript reverse transcriptase. After ligation of
      EcoRI adapters (5'-AATTCGACGAGG-3') followed by kinasing
      adapters and by XhoI digestion, the cDNA was size selected
      by chromatography on Sepharose CL-2B columns and fractions
      containing cDNAs larger than 1000 bp were ligated into
      EcoRI/XhoI-digested pCS107. Reference for library
      construction: Current Genomics 4, 635-644. Library
      constructed by Michelle Tabb and Bruce Blumberg (Dept of
      Developmental and Cell Biology, University of California,
      Irvine)."
  Query Match 54.3%; Score 22.8; DB 9; Length 864;
  Best Local Similarity 77.1%; Pred.No. 2.1e+02;
  Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 57
AA615518
LOCUS
DEFINITION
  v070e02.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
  IMAGE:1064474 5', mRNA sequence.
ACCESSION
AA615518
VERSION
KEYWORDS
SOURCE
  AA615518.1 GI:2502746
  EST.
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muridea; Muridae; Murinae; Mus.
  1 (bases 1 to 454)
  Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
  Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
  Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
  Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
  Waterston,R.
  The WashU-HMNI Mouse EST Project
  Unpublished (1996)
  Contact: Marra M/Mouse EST Project
  WashU-HMNI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:586834
  Seq primer: -28ml3 rev2 ET from Amersham.
  Location/Qualifiers
    1..454
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C3H"
      /db_xref="taxon:10090"
      /clone="IMAGE:1064474"
      /cell_line="C2C12"
      /lab_host="DH10B"
      /clone_lib="Barstead mouse myotubes MPLRB5"
      /note="Vector: pT7T3D-PacI; Site_1: EcoRI; Site_2: NotI;
      1st strand cDNA was primed with a Not I - oligo(dT) primer
      [5.
      TGTACGAATCTGAAGTGGAGGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
      3']; double-stranded cDNA was ligated to Eco RI adaptors
      [AATTCGACGCTTGT], digested with Not I and cloned into the
      Not I and Eco RI sites of the modified pT7T3 vector.
      Library constructed by Bob Barstead. The C2C12 cell line
      (available from ATCC, catalog # CRL-1772) differentiates
      rapidly, forming contractile myotubes and producing
      characteristic muscle proteins.
  Query Match 53.8%; Score 22.6; DB 1; Length 454;
  Best Local Similarity 75.7%; Pred.No. 2.3e+02;
  Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

RESULT 58
CB818826
LOCUS
DEFINITION
  1 GCCCAGCGTTCGCCGAGCGCCCTTACAAAGACTATGAAG 37
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  326 GCCCAGCGCGCCCATGTCCGTACATGACTACGAG 362
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Query Match 54.3%; Score 22.8; DB 9; Length 864;
  Best Local Similarity 77.1%; Pred.No. 2.1e+02;
  Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy 8 CTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 42
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 820 CTTGGACGACATTTACAAAGCATATNCAGTAAAG 854

RESULT 57
AA615518
LOCUS
DEFINITION
  v070e02.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
  IMAGE:1064474 5', mRNA sequence.
ACCESSION
AA615518
VERSION
KEYWORDS
SOURCE
  AA615518.1 GI:2502746
  EST.
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muridea; Muridae; Murinae; Mus.
  1 (bases 1 to 454)
  Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
  Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
  Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
  Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
  Waterston,R.
  The WashU-HMNI Mouse EST Project
  Unpublished (1996)
  Contact: Marra M/Mouse EST Project
  WashU-HMNI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:586834
  Seq primer: -28ml3 rev2 ET from Amersham.
  Location/Qualifiers
    1..454
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C3H"
      /db_xref="taxon:10090"
      /clone="IMAGE:1064474"
      /cell_line="C2C12"
      /lab_host="DH10B"
      /clone_lib="Barstead mouse myotubes MPLRB5"
      /note="Vector: pT7T3D-PacI; Site_1: EcoRI; Site_2: NotI;
      1st strand cDNA was primed with a Not I - oligo(dT) primer
      [5.
      TGTACGAATCTGAAGTGGAGGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
      3']; double-stranded cDNA was ligated to Eco RI adaptors
      [AATTCGACGCTTGT], digested with Not I and cloned into the
      Not I and Eco RI sites of the modified pT7T3 vector.
      Library constructed by Bob Barstead. The C2C12 cell line
      (available from ATCC, catalog # CRL-1772) differentiates
      rapidly, forming contractile myotubes and producing
      characteristic muscle proteins.
  Query Match 53.8%; Score 22.6; DB 1; Length 454;
  Best Local Similarity 75.7%; Pred.No. 2.3e+02;
  Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

RESULT 58
CB818826
LOCUS
DEFINITION
  1 GCCCAGCGTTCGCCGAGCGCCCTTACAAAGACTATGAAG 37
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  326 GCCCAGCGCGCCCATGTCCGTACATGACTACGAG 362
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Query Match 54.3%; Score 22.8; DB 9; Length 864;
  Best Local Similarity 77.1%; Pred.No. 2.1e+02;
  Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

EST 1716 Green Apricot Fruit Lambda Zap II Library Prunus armeniaca
cDNA clone ba003d02 3', mRNA sequence.

CB818826
CB818826.1 GI:40715236
EST.

Prunus armeniaca (apricot)
Prunus armeniaca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.

REFERENCE

1 (bases 1 to 754)
Grimplet, J., Romieu, C., Audergon, J.M., Albagnac, G., Lambert, P.,
Boucher, J.P. and Terrier, N.

High Throughput Detection of Isogenes among 5724 3' EST from
Apricot Fruit (Prunus armeniaca)

Unpublished (2003)

Contact: Audergon JM

Unité de génétique et amélioration des fruits et légumes

Institut National de la Recherche Agronomique

Domaine Saint-Maurice BP 94 84143 Montfavet cedex

Tel: 00-33-(0)4-32-72-26-68

Fax: 00-33-(0)4-32-72-26-62

Email: audergon@avignon.inra.fr

Seq primer: 17.

FEATURES

source
Location/Qualifiers
1..754
/organism="Prunus armeniaca"
/mol_type="mRNA"
/cultivar="Bergeron"
/db_xref="taxon:36596"
/clone="ba003d02"
/dev_stage="Green stage"
/clone_lib="Green Apricot Fruit Lambda Zap II Library"
/note="Organ: Fruit; Vector: Lambda Zap II; Site_1: Eco
RI; Site_2: XhoI; Oriented library, construction described
in Molecular cloning and expression of a cDNA encoding
1-aminocyclopropane-1-carboxylate (ACC) oxidase from
apricot fruit (Prunus armeniaca cv. Bergeron) by
Mbequie-Mbequie D, Chahine H, Gomez RM, Gouble B, Audergon
JM, Soury W, Albagnac G, Flis-Lyaon B in Physiol Plant
105:294-303 1999"

ORIGIN

Query Match 53.8%; Score 22.6; DB 4; Length 754;
Best Local Similarity 75.7%; Pred. No. 2.5e+02;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 4 CACGCTGGCCGAGCCCTTACAAAGACTATGAAGTAA 40
|||||
Db 89 CACCATGGCGCGTCCGTACAAAGACCGTGAAGAAA 125
|||||

RESULT 59
DV624038
LOCUS
DEFINITION
93529.1 Cold Sweetening C Solanum tuberosum cDNA clone 93529 5',
mRNA sequence.

ACCESSION
DV624038

VERSION
DV624038.1 GI:78747635

KEYWORDS
EST.

SOURCE
Solanum tuberosum (potato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 609)

REFERENCE

1
Flinn, B., Rothwell, C., Sardana, R., Griffiths, R., Lague, M., De
Koeyer, D., Audy, P., Gover, C., Li, X-Q., Wang-Pruski, G. and Regan, S.

Generation of ESTs from tubers following 3 months storage at 4
degrees Celsius, and 95% relative humidity, then 3 days of
reconditioning at 21 degrees

JOURNAL

Unpublished (2005)

Contact: Vicki Gustafson

FEATURES

source
Location/Qualifiers
1..609
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Shepody"
/db_xref="taxon:4113"
/clone="93529"
/tissue_type="Tubers"
/lab_host="XL10-Gold"
/clone_lib="Cold Sweetening C"
/note="Vector: pBluescript II SK(+) XR; Site_1: EcoRI;
Site_2: XhoI; supplier: Developmental series. Plants from
pathogen-free Solanum tuberosum var. Shepody, clone 1756,
nuclear stock were grown in a greenhouse under natural
conditions. Mature, harvested tubers were stored in the
dark at 4C, 95% relative humidity for 3 months, and then
reconditioned at 21C for three days. RNA was isolated for
library construction."

ORIGIN

Query Match 53.3%; Score 22.4; DB 10; Length 609;
Best Local Similarity 72.5%; Pred. No. 2.9e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 CCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 41
|||||
Db 149 CCATGCTTGCACGACCATGAACAAGATTATGAAGTCAA 188
|||||

RESULT 60

AQ855455/c

LOCUS

DEFINITION

CPG1823A CpIOWAGD1 Cryptosporidium parvum genomic, genomic survey

sequence.

ACCESSION

AQ855455

VERSION

AQ855455.1 GI:6205912

KEYWORDS

GSS.

SOURCE

Cryptosporidium parvum

Cryptosporidium parvum

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Cryptosporidiidae; Cryptosporidium.

1 (bases 1 to 746)

Strong, W.B. and Nelson, R.G.

Preliminary profile of the Cryptosporidium parvum genome: an

expressed sequence tag and genome survey sequence analysis

Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)

10717299

Contact: Nelson, R. G.

Depts. of Medicine & Pharmaceutical Chemistry

San Francisco General Hospital-University of California, San

Francisco

Box 0811, San Francisco, CA 94143-0811, USA

Tel: 415 206 8846

Fax: 415 206 3353

Email: malaria@itsa.ucsf.edu

For Annotation Data see http://medsfgh.ucsf.edu/id/CpTags/home.html

Seq primer: M13(-21) forward

Class: shotgun.

FEATURES

source

Location/Qualifiers

1..746

/organism="Cryptosporidium parvum"

/mol_type="genomic DNA"

/strain="IOWA"

/db_xref="taxon:5807"

/laE_host="E. coli XL2 Blue MRF"

/clone_lib="CpIOWAGD1"

/note="vector: pBluescript II (SK-); Site_1: EcoRV; C.

parvum (IOWA isolate) genomic DNA was hydrodynamically

sheared to produce fragments having a tight size

The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: ygunstafson@bioatlantech.nb.ca
Seq primer: 13.

distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center (<http://sequence-www.stanford.edu/group/techdev/shear.htm>). The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoRI V-digested, alkaline phosphatase-treated pBluescript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF+. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match 53.3%; Score 22.4; DB 11; Length 746;
Best Local Similarity 72.5%; Pred. No. 3e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 695 CCAGGCTTTCTAAGCTCTTCAAATATTTATGAAGGAAG 656

RESULT 61

CK718467 765 bp mRNA linear EST 20-OCT-2005
LOCUS 18526 Swollen Stolon Solanum tuberosum cDNA, mRNA sequence.
DEFINITION CK718467
VERSION CK718467.1 GI:42510181
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 765)
Flinn, B., Rothwell, C., Sardana, R., Griffiths, R., Lague, M., DeKoeyer, D., Audy, P., Goyer, C., Li, X.-Q., Wang-Pruski, G. and Regan, S.
Generation of ESTs from swollen stolon tissues of potato
Contact: Vicki Gustafson
Unpublished (2004)
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: vgustafson@bioatlantech.nb.ca
bf1inn@bioatlantech.nb.ca
Seq primer: T3.

REFERENCE

AUTHORS

Flinn, B., Rothwell, C., Sardana, R., Griffiths, R., Lague, M., DeKoeyer, D., Audy, P., Goyer, C., Li, X.-Q., Wang-Pruski, G. and Regan, S.

TITLE

JOURNAL

COMMENT

The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: vgustafson@bioatlantech.nb.ca
bf1inn@bioatlantech.nb.ca
Seq primer: T3.

FEATURES

source

1..765
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Shepody"
/db_xref="taxon:4113"
/tissue_type="Stolon"
/lab_hosts="XL10-Gold"

/clone_libs="Swollen Stolon"
/note="Vector: pBluescript II SK(+); Site 1: EcoRI; Site 2: XhoI; supplier: Developmental series. Plants from pathogen-free Solanum tuberosum var. Shepody, clone 1756, nuclear stock were grown in a screenhouse under natural conditions. RNA was isolated from swollen stolon tissue, 3-10mm in length, which was cut from the tip, to the base of swelling."

ORIGIN

Query Match 53.3%; Score 22.4; DB 5; Length 765;
Best Local Similarity 72.5%; Pred. No. 3e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 41
Db 620 CCATTGCTTGCCTCAAGCCATGACAAAGATTATGAAGTCAA 659

RESULT 62

LOCUS

DEFINITION

CV286484 810 bp mRNA linear EST 20-OCT-2005

66768.1 After-Cooking Darkening A Solanum tuberosum cDNA clone

66768.5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2004)
Contact: Vicki Gustafson
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: vgustafson@bioatlantech.nb.ca
Seq primer: T3.

FEATURES

source

1..810
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Shepody"
/db_xref="taxon:4113"
/clone_libs="After-Cooking Darkening A"
/note="Vector: pBluescript II SK(+); Site 1: EcoRI; Site 2: XhoI; supplier: Developmental series. Plants from pathogen-free Solanum tuberosum var. Shepody, clone 1756, nuclear stock were grown in a screenhouse under natural conditions. Mature, harvested tubers were stored in the dark at 9C, 95% relative humidity for 1 month. The tuber skin was then removed, and RNA was isolated from the skinless tubers and used for library construction."

ORIGIN

Query Match 53.3%; Score 22.4; DB 8; Length 810;
Best Local Similarity 72.5%; Pred. No. 3.1e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 41

Db 472 CCGTGTGCTGCCCAAGCCATGACAAAGATTATGAAGTCAA 511

RESULT 63

LOCUS

DEFINITION

BE004395 338 bp mRNA linear EST 05-JUN-2000

CM0-EN0105-180300-295-e05 BN0105 Homo sapiens cDNA, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 338)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE

JOURNAL

COMMENT

Unpublished (2004)
Contact: Vicki Gustafson
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: vgustafson@bioatlantech.nb.ca
Seq primer: T3.

FEATURES

source

1..338
/organism="Homo sapiens"
/mol_type="mRNA"
/cultivar="Shepody"
/db_xref="taxon:4113"
/tissue_type="Stolon"
/lab_hosts="XL10-Gold"

/clone_libs="Swollen Stolon"
/note="Vector: pBluescript II SK(+); Site 1: EcoRI; Site 2: XhoI; supplier: Developmental series. Plants from pathogen-free Solanum tuberosum var. Shepody, clone 1756, nuclear stock were grown in a screenhouse under natural conditions. RNA was isolated from swollen stolon tissue, 3-10mm in length, which was cut from the tip, to the base of swelling."

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=CMO-BN0105-180
300-295-e05&t3=2000-03-18&t4=1)

Seq primer: puc 18 forward
High quality sequence stop: 336.

FEATURES

source

```

1..338
   Location/Qualifiers
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /dev_stage="Adult"
     /clone_lib="BN0105"
     /note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

ORIGIN

```

Query Match      52.9%; Score 22.2; DB 7; Length 338;
Best Local Similarity 77.1%; Pred. No. 3.2e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAA 41
    |||||  |||||  |||||  |||||  |||||  |||||
Db 231 GGTGGCCCACTCTCACAAATACAAAGAGTAAA 265.

```

```

RESULT 64
AA956676
LOCUS      AA956676               362 bp      mRNA      linear      EST 04-JUL-1999
DEFINITION UI-R-E1-fl-g-03-0-UI.s1 UI-R-E1 Rattus norvegicus cDNA clone
            UI-R-E1-fl-g-03-0-UI 3', similar to gi|2888286|gb|AA818700|AA818700
            UI-R-A0-az-c-06-0-UI.s1 UI-R-A0 Rattus norvegicus cDNA clone
            UI-R-A0-az-c-06-0-UI 3', similar to gi|1808536|gb|AA210199|AA210199
            mu40c01.r1 Soares 2NbMT Mus musculus cDNA clone 641856 5', mRNA
            sequence.

```

```

ACCESSION  AA956676
VERSION     AA956676.1   GI:4239150
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
            1 (bases 1 to 362)
REFERENCE   Bonaldo,M.F., Lennon,G. and Soares,M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
JOURNAL     On May 7, 1998 this sequence version replaced gi:3120371.
PUBMED     8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565

```

FEATURES

source

```

1..362
   Location/Qualifiers
     /organism="Rattus norvegicus"
     /mol_type="mRNA"
     /strain="Sprague-Dawley"
     /db_xref="taxon:10116"
     /clone="UI-R-E1-fl-g-03-0-UI"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies)"
     /clone_lib="UI-R-E1"
     /note="Vector: pYT73D-PacI; Site 1: Not I; Site 2: Eco RI; The UI-R-E1 library is a subtracted library derived from a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-E1) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-E0 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-E0 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-E1 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."

```

ORIGIN

```

Query Match      52.9%; Score 22.2; DB 1; Length 362;
Best Local Similarity 88.9%; Pred. No. 3.2e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    |||||  |||||  |||||  |||||  |||||  |||||
Db 270 GAGCCCTTACAAAGACTATGAAGTAAA 296

```

```

RESULT 65
CB727988
LOCUS      CB727988               482 bp      mRNA      linear      EST 11-APR-2003
DEFINITION ANGNNUC:URRG1-00029-C2-A urrg1 (14046) Rattus norvegicus cDNA clone
            urrg1-00029-c2 5', mRNA sequence.

```

```

ACCESSION  CB727988
VERSION     CB727988.1   GI:29795061
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus

```

```

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
            1 (bases 1 to 482)
AUTHORS     Angen EST Program.
TITLE       Angen Rat EST Program
JOURNAL     Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Angen, Inc
            One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881

```

Email: bento-soares@uiowa.edu

The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized adult 8-Day-Embryo library. cDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics This clone is also available through the I.M.A.G.E. Consortium at LBNL (info@image.lbnl.gov). IMAGE ID=1772096 The following repetitive elements were found in this cDNA sequence: 1-37, >AT_richLow_complexity

Seq primer: M13 Forward

PolyA=No.

```
FEATURES
source
    Plate: 00029 row: c column: 2.
    Location/Qualifiers
      1..482
        /organism="Rattus norvegicus"
        /mol_type="mRNA"
        /db_xref="taxon:10116"
        /clone="urrgi-00029-c2"
        /clone_lib="urrgi (14046)"
        /note="Vector: pT7T3D-PacI"
ORIGIN
    Query Match      52.9%; Score 22.2; DB 4; Length 482;
    Best Local Similarity 88.9%; Pred. No. 3.4e+02;
    Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    |||||
Db 403 GAGCCCTTACAAAGACTATGAAGCACA 429
    |||||
RESULT 66
LOCUS A1171672 484 bp mRNA linear EST 20-JAN-1999
DEFINITION EST217647 Normalized rat muscle, Bento Soares Rattus sp. cDNA clone
ACCESSION RMUB014.3, end, mRNA sequence.
VERSION A1171672
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
    1 (bases 1 to 484)
    Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
    Kerlavage, A.R. and Adams, M.D.
    Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
    Gene Index
    Unpublished (1998)
    Contact: Lee, NH
    The Institute for Genomic Research
    9712, Medical Center Drive, Rockville, MD 20850, USA
    Tel: (301)-838-3529
    Fax: (301)-838-0208
    Email: nhlee@tigr.org
    Seq primer: M13-21.
    Site_2: NotI"
FEATURES
source
    Location/Qualifiers
      1..484
        /organism="Rattus sp."
        /mol_type="mRNA"
        /db_xref="ATCC (inhost):2028324"
        /db_xref="taxon:10118"
        /clone="RMUB014"
        /clone_lib="Normalized rat muscle, Bento Soares"
        /note="Organ: muscle; Vector: pT7T3Pac; Site_1: EcoRI;
        Site_2: NotI"
ORIGIN
    Query Match      52.9%; Score 22.2; DB 1; Length 484;
    Best Local Similarity 88.9%; Pred. No. 3.4e+02;
    Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    |||||
Db 356 GAGCCCTTACAAAGACTATGAAGCACA 382
    |||||
RESULT 67
LOCUS BP470636/c 498 bp mRNA linear EST 26-JAN-2006
DEFINITION BP470636 Rattus norvegicus pancreatic islet Rattus norvegicus cDNA
clone RBC06466 3', mRNA sequence.
ACCESSION BP470636
```

```
VERSION BP470636.1 GI:85732193
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
    1 (bases 1 to 498)
    Takeda, J., Wang, H. and Horikawa, Y.
    Expression profile of mRNAs from rat pancreatic islet
    Unpublished (2006)
    Contact: Yukio Horikawa
    Laboratory of Molecular Genetics
    Institute for Molecular and Cellular Regulation, Gunma University
    Showa-machi 3-39-15, Maebashi, Gunma, 371-8512, Japan
    Tel: 81-27-220-8832
    Fax: 81-27-220-8889
    Email: yhorikaw@showa.gunma-u.ac.jp.
FEATURES
source
    Location/Qualifiers
      1..498
        /organism="Rattus norvegicus"
        /mol_type="mRNA"
        /db_xref="taxon:10116"
        /clone="RBC06466"
        /tissue type="pancreatic islet"
        /clone_lib="Rattus norvegicus pancreatic islet"
ORIGIN
    Query Match      52.9%; Score 22.2; DB 3; Length 498;
    Best Local Similarity 88.9%; Pred. No. 3.4e+02;
    Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    |||||
Db 420 GAGCCCTTACAAAGACTATGAAGCACA 394
    |||||
RESULT 68
LOCUS A1175079 515 bp mRNA linear EST 20-JAN-1999
DEFINITION EST218597 Normalized rat muscle, Bento Soares Rattus sp. cDNA clone
RMUCCL15.3, end, mRNA sequence.
ACCESSION A1175079
VERSION A1175079.1 GI:3725717
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
    1 (bases 1 to 515)
    Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
    Kerlavage, A.R. and Adams, M.D.
    Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
    Gene Index
    Unpublished (1998)
    Contact: Lee, NH
    The Institute for Genomic Research
    9712, Medical Center Drive, Rockville, MD 20850, USA
    Tel: (301)-838-3529
    Fax: (301)-838-0208
    Email: nhlee@tigr.org
    Seq primer: M13-21.
    Location/Qualifiers
      1..515
        /organism="Rattus sp."
        /mol_type="mRNA"
        /db_xref="ATCC (inhost):2032225"
        /db_xref="taxon:10118"
        /clone="RMUCCL15"
        /clone_lib="Normalized rat muscle, Bento Soares"
        /note="Organ: muscle; Vector: pT7T3Pac; Site_1: EcoRI;
        Site_2: NotI"
FEATURES
source
    Location/Qualifiers
      1..515
        /organism="Rattus sp."
        /mol_type="mRNA"
        /db_xref="ATCC (inhost):2032225"
        /db_xref="taxon:10118"
        /clone="RMUCCL15"
        /clone_lib="Normalized rat muscle, Bento Soares"
        /note="Organ: muscle; Vector: pT7T3Pac; Site_1: EcoRI;
        Site_2: NotI"
```



```

AA818700      584 bp      mRNA      linear      EST 05-FEB-1999
LOCUS         UI-R-A0-az-c-06-0-UI.s1 UI-R-A0 Rattus norvegicus cDNA clone
DEFINITION    UI-R-A0-az-c-06-0-UI 3', similar to gi|1808536|gb|AA210199|AA210199
               mu40c01.r1 Soares 2NbMT Mus musculus cDNA clone 641856 5', mRNA
               sequence.
ACCESSION     AA818700      GI:4227382
VERSION       AA818700.1
KEYWORDS      Rattus norvegicus (Norway rat)
SOURCE        EST.
ORGANISM      Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE     1 (bases 1 to 584)
AUTHORS       Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE         Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL       Genome Res. 6 (9), 791-806 (1996)
PUBMED        8889548
COMMENT       On Feb 17, 1998 this sequence version replaced gi.2888286.
               Contact: Soares, MB
               Coordinated Laboratory for Computational Genomics
               University of Iowa
               375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
               Tel: 319 335 8250
               Fax: 319 335 9565
               Email: bento-soares@uiowa.edu
               The sequence tag present in the cDNA between the NotI site and the
               oligo-dT track served to identify it as a clone from the normalized
               adult lung library. cDNA Library Preparation: M. Fatima Bonaldo,
               Ph.D. Clone distribution: clones will be available through Research
               Genetics
               Seq primer: M13 Forward.
               Location/Qualifiers
               1..584
               /organism="Rattus norvegicus"
               /mol_type="mRNA"
               /strain="Sprague-Dawley"
               /db_xref="taxon:10116"
               /clone="UI-R-A0-az-c-06-0-UI"
               /dev_stage="adult"
               /lab_host="DH10B (Life Technologies)"
               /clone_lib="UI-R-A0"
               /note="Vector: pRT73D-PacI; Site1: Not I; Site2: Eco RI;
               This library consists of a mixture of individually tagged
               normalized libraries constructed from rat placenta, adult
               lung, brain, liver, kidney, heart, spleen, ovary, and
               muscle. The tag is a string of 3-5 nucleotides present
               between the Not I site and the oligo-dT track which allows
               identification of the library of origin of a clone within
               the mixture."

ORIGIN
Query Match      52.9%; Score 22.2; DB 1; Length 584;
Best Local Similarity 88.9%; Pred. No. 3.5e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      15 GAGCCCTTACAAAGACTATGAAGTAAA 41
        |||||
Db      373 GAGCCCTTACAAAGACTATGAAGTAAA 399

RESULT 75
LOCUS      AI599233/c
DEFINITION EST250936 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
ACCESSION  AI599233
VERSION     AI599233.1 GI:4608281
KEYWORDS    Rattus sp.
SOURCE      EST.
ORGANISM    Rattus sp.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE     1 (bases 1 to 612)
AUTHORS       Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
               Kerlavage,A.R. and Adams,M.D.
TITLE         Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
               Gene Index
JOURNAL       Unpublished (1998)
PUBMED
COMMENT       Other ESTs: TC74592
               Contact: Lee, NH
               The Institute for Genomic Research
               9712, Medical Center Drive, Rockville, MD 20850, USA
               Tel: (301)-838-3529
               Fax: (301)-838-0208
               Email: nhlee@tigr.org
               Seq primer: M13-21.
               Location/Qualifiers
               1..612
               /organism="Rattus sp."
               /mol_type="mRNA"
               /db_xref="taxon:10118"
               /clone="REMEC64"
               /dev_stage="embryo 8, 12, 18 dpc"
               /clone_lib="Normalized rat embryo, Bento Soares"
               /note="Vector: pRT73Pac; Site1: EcoRI; Site2: NotI"

ORIGIN
Query Match      52.9%; Score 22.2; DB 1; Length 612;
Best Local Similarity 88.9%; Pred. No. 3.5e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      15 GAGCCCTTACAAAGACTATGAAGTAAA 41
        |||||
Db      419 GAGCCCTTACAAAGACTATGAAGTAAA 393

RESULT 76
LOCUS      BM386471
DEFINITION UI-R-CN1-cjg-g-03-0-UI.s1 UI-R-CN1 Rattus norvegicus cDNA clone
ACCESSION   BM386471
VERSION     BM386471.1 GI:18186524
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE     1 (bases 1 to 614)
AUTHORS       Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE         Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL       Genome Res. 6 (9), 791-806 (1996)
PUBMED        8889548
COMMENT       Contact: Soares, MB
               Coordinated Laboratory for Computational Genomics
               University of Iowa
               375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
               Tel: 319 335 8250
               Fax: 319 335 9565
               Email: bento-soares@uiowa.edu
               The sequence contained an oligo-dT track that was present in the
               oligonucleotide that was used to prime the synthesis of first
               strand cDNA and therefore this may represent a bonafide poly A
               tail. The sequence tag present in the cDNA between the NotI site
               and the oligo-dT track served to identify it as a clone from the
               normalized cervix library cDNA Library Preparation: M.B. Soares Lab
               Clone distribution: clones will be available through Research
               Genetics (www.resgen.com) The following repetitive elements were
               found in this cDNA sequence: 1-30, >POLY_A$imple_repeat
               Seq primer: M13 Forward

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE     1 (bases 1 to 612)
AUTHORS       Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
               Kerlavage,A.R. and Adams,M.D.
TITLE         Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
               Gene Index
JOURNAL       Unpublished (1998)
PUBMED
COMMENT       Other ESTs: TC74592
               Contact: Lee, NH
               The Institute for Genomic Research
               9712, Medical Center Drive, Rockville, MD 20850, USA
               Tel: (301)-838-3529
               Fax: (301)-838-0208
               Email: nhlee@tigr.org
               Seq primer: M13-21.
               Location/Qualifiers
               1..612
               /organism="Rattus sp."
               /mol_type="mRNA"
               /db_xref="taxon:10118"
               /clone="REMEC64"
               /dev_stage="embryo 8, 12, 18 dpc"
               /clone_lib="Normalized rat embryo, Bento Soares"
               /note="Vector: pRT73Pac; Site1: EcoRI; Site2: NotI"

ORIGIN
Query Match      52.9%; Score 22.2; DB 1; Length 612;
Best Local Similarity 88.9%; Pred. No. 3.5e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      15 GAGCCCTTACAAAGACTATGAAGTAAA 41
        |||||
Db      419 GAGCCCTTACAAAGACTATGAAGTAAA 393

RESULT 76
LOCUS      BM386471
DEFINITION UI-R-CN1-cjg-g-03-0-UI.s1 UI-R-CN1 Rattus norvegicus cDNA clone
ACCESSION   BM386471
VERSION     BM386471.1 GI:18186524
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE     1 (bases 1 to 614)
AUTHORS       Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE         Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL       Genome Res. 6 (9), 791-806 (1996)
PUBMED        8889548
COMMENT       Contact: Soares, MB
               Coordinated Laboratory for Computational Genomics
               University of Iowa
               375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
               Tel: 319 335 8250
               Fax: 319 335 9565
               Email: bento-soares@uiowa.edu
               The sequence contained an oligo-dT track that was present in the
               oligonucleotide that was used to prime the synthesis of first
               strand cDNA and therefore this may represent a bonafide poly A
               tail. The sequence tag present in the cDNA between the NotI site
               and the oligo-dT track served to identify it as a clone from the
               normalized cervix library cDNA Library Preparation: M.B. Soares Lab
               Clone distribution: clones will be available through Research
               Genetics (www.resgen.com) The following repetitive elements were
               found in this cDNA sequence: 1-30, >POLY_A$imple_repeat
               Seq primer: M13 Forward

```

```

POLYA=Yes.
FEATURES
    source
        Location/Qualifiers
            1..614
                /organism="Rattus norvegicus"
                /mol_type="mRNA"
                /strain="Sprague-Dawley"
                /db_xref="taxon:10116"
                /clone="UI-R-CNI-cjg-g-03-0-UI"
                /dev_stage="adult"
                /lab_host="DH10B (Life Technologies)"
                /clone_lib="UI-R-CNI"
                /note="Vector: p773D-PacI; Site 1: Not I; Site 2: Eco RI;
The UI-R-CNI library is a subtracted library derived from
the following pool of seven normalized rat libraries:
normalized rat seminal vesicles, normalized rat penis,
normalized rat bladder, normalized rat cervix, normalized
rat brown adipose, normalized rat fundus, and normalized
rat salivary gland. It was constructed according to the
procedure described by Bonaldo, Lennon & Soares (Genome
Research Genome 6: 791-806, 1996). For construction of
the CNI library, plasmid DNA from the pool of normalized
libraries was electroporated into competent bacteria for
the production of single-stranded circular DNA. This was
then used as a tracer in a subtractive hybridization with
a driver (PCR amplified inserts from a plamid DNA template
preparation) comprising: a) a pool of about 34,000 clones
from the Rat Unigene Set corresponding to plates R-5-AA-NN
excluding plates R-5-MM and MN. This pool represented 40%
of the final driver population. b) a pool of about 29,000
clones from subtracted libraries CA0 and CA1 corresponding
to plates R-CA0-AMV through R-CA0-AXS, R-CA0-AXZ through
R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS,
R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through
R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA,
R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through
R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through
R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BUR,
R-CA1-BVT through R-CA1-BKE, R-CA1-BXJ, R-CA1-BKF,
R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through
R-CA1-BLN, R-CA1-BLS, R-CA1-BLU-V, R-CA1-BNR, and
R-CA1-BLE. The resulting pool represented 20% of the
final driver population. c) a pool of about 15,000 clones
from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s
and normalized libraries CS0, CT0, CU0, CW0, and CX0
corresponding to plates R-CS0s-CBD through R-CS0s-CBO,
R-CT0s-CCA through R-CT0s-CAX, R-CU0s-CBP through
R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN
through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through
R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP,
R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through
R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through
R-CX0-BXM. The resulting pool represented 5% of the final
driver population. d) a pool of about 5,000 clones (1,000
from non-normalized eye library CV0 and 4,000 from
normalized eye library CV1) corresponding to plates
R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC,
R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through
R-CV1-BVU. This pool represented about 5% of the final
driver population. e) A pool of about 10,000 clones from
subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA
library fraction from rat whole embryo), and BX0 (0.5-7kb
cDNA library fraction from rat whole embryo) corresponding
to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through
R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AOY
through R-BX0-ASH. The resulting pool represented 5% of
the final driver population. f) a pool of about 7,000
clones from the seven non-normalized libraries that make
up the tracer including CY0, CZ0, DA0, DB0, DC0, DD0, and
DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ,
R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ
through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ
through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY
through R-DC0-CBA, R-DD0-BZR through R-DD0-CAA,

```

R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLM-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CNO pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-b-10-0-UI, bkw-d-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-h-12-0-UI, bky-a-05-0-UI, bkr-a-06-0-UI, bkr-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blg-h-04-0-UI, blr-a-03-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CNO clone (corresponding to the address bkr-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG TISSUE=cervix
TAG LIB=UI-R-CNI
TAG_SEQ=GACCA"

ORIGIN

Query Match	52.9%	Score 22.2;	DB 2;	Length 614;
Best Local Similarity	88.9%	Pred. No. 3.5e+02;		
Matches 24;	Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
|||||
Db 386 GAGCCCTTACAAAGACTATGAAGTAAA 412
|||||

RESULT 77
AA946356 620 bp mRNA linear EST 16-JUN-1998
LOCUS EST201855 Normalized rat lung, Bento Soares Rattus sp. cDNA clone
DEFINITION RLBH16 3' end, mRNA sequence.
AA946356
ACCESSION AA946356.1 GI:3106272
VERSION EST.
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 620)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat
Gene Index
Unpublished (1998)
CONTACT: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igr.org
Seq primer: M13-21.
Location/Qualifiers
1..620
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC (inhost):2015290"
/db_xref="taxon:10118"
/clone="RLBH16"
/clone_lib="Normalized rat lung, Bento Soares"

/note="Organ: lung; Vector: pT7T3pac; Site_1: EcoRI;
Site_2: NotI"

ORIGIN

Query Match 52.9%; Score 22.2; DB 1; Length 620;
Best Local Similarity 88.9%; Pred. No. 3.5e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
|||||
Db 338 GAGCCCTTACAAAGACTAAGAACCAAA 364
|||||

RESULT 78

CB326875 CB326875 673 bp mRNA linear EST 12-AUG-2004
LOCUS UI-R-F50-crw-i-16-0-UI.s1 NCI_CGAP_F50 Rattus norvegicus cDNA clone
DEFINITION IMAGE:7360314 3', mRNA sequence.

ACCESSION

CB326875

VERSION

CB326875.1 GI:28861533

KEYWORDS

EST

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 673)

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

JOURNAL

discovery

PUBMED

889548

COMMENT

Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
IMAGE (http://image.llnl.gov)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1..673 Location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clones="IMAGE:7360314"
/dev_stage="ADULT"
/lab_hosts="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_F50"

/note="Vector: pT7T3D-PacI; Site 1: Not I; Site 2: Eco RI;
UI-R-F50 is a cDNA library constructed from Swarm Rat
Chondrosarcoma. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGCCGCCGAT. The tumor was
provided by Dr Jeff Stevens from University of Iowa.
TAG_TISSUE=Rat SRC-JWS Cell Line
TAG_LIB=UI-R-F50
TAG_SEQ=AGCCGCCGAT"

ORIGIN

Query Match 52.9%; Score 22.2; DB 4; Length 673;
Best Local Similarity 88.9%; Pred. No. 3.6e+02;

Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
|||||
Db 374 GAGCCCTTACAAAGACTAAGAACCAAA 400
|||||

RESULT 79

CA504656 CA504656 689 bp mRNA linear EST 14-NOV-2002
LOCUS UI-R-FJ0-cpw-h-08-0-UI.s1 UI-R-FJ0 Rattus norvegicus cDNA clone
DEFINITION UI-R-FJ0-cpw-h-08-0-UI 3', mRNA sequence.

ACCESSION

CA504656

VERSION

CA504656.1 GI:24995610

KEYWORDS

EST

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 689)

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

JOURNAL

discovery

PUBMED

889548

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-30, >POLY A#Simple_repeat (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1..689 Location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clones="UI-R-FJ0-cpw-h-08-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_hosts="DH10B (Life Technologies)"
/clone_lib="UI-R-FJ0"

/note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
UI-R-FJ0 is a cDNA library containing the following
tissues): rat embryo. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CATCTCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant)
TAG_TISSUE=rat-embryo
TAG_LIB=UI-R-FJ0
TAG_SEQ=CATCTCTACT"

ORIGIN

```

Query Match      52.9%; Score 22.2; DB 4; Length 689;
Best Local Similarity 88.9%; Pred. No. 3.6e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    ||||| ||||| ||||| ||||| |||||
Db 386 GAGCCCTTACAAAGACTATGAAGACCAA 412

RESULT 80
LOCUS CK838771 736 bp mRNA linear EST 05-MAR-2004
DEFINITION UI-R-AC1-xx-c-10-0-UI.s10 UI-R-AC1 Rattus norvegicus cDNA clone
            UI-R-AC1-xx-c-10-0-UI 3', mRNA sequence.
ACCESSION CK838771
VERSION CK838771.1 GI:45189141
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
1 (bases 1 to 736)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/rat.html
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES             source
1..736
    /location=Qualifiers
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /strain="Sprague-Dawley"
    /db_xref="taxon:10116"
    /clones="UI-R-AC1-xx-c-10-0-UI"
    /dev_stages="adult"
    /lab_host="DH10B (Life Technologies)"
    /clone_lib="UI-R-AC1"
    /note="Vector: pT73D-PacI; Site 1: Not 1; Site 2: Eco RI;
The UI-R-AC1 library is a normalized library constructed
from 16.5 dpc rat atrioventricular (AV) canal. The tag
is a string of 5 nucleotides present between the Not I
site and the oligo-dT track. The library was constructed
as described byonaldo, Lennon and Soares, Genome
Research 6: 791-806, 1996. Tissue provided by Jim Lin,
Department of Biology, University of Iowa.
TAG_TISSUE=muscle
TAG_LIB=UI-R-AC1
TAG_SEQ=AAG"

ORIGIN
Query Match      52.9%; Score 22.2; DB 5; Length 736;
Best Local Similarity 88.9%; Pred. No. 3.6e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    ||||| ||||| ||||| ||||| |||||
Db 374 GAGCCCTTACAAAGACTATGAAGACCAA 400

Query Match      52.9%; Score 22.2; DB 5; Length 736;
Best Local Similarity 88.9%; Pred. No. 3.6e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAAA 41
    ||||| ||||| ||||| ||||| |||||
Db 374 GAGCCCTTACAAAGACTATGAAGACCAA 400

```

```

RESULT 81
LOCUS CK2983576 859 bp DNA linear GSS 11-AUG-2005
DEFINITION 197310 Tomato MboI BAC Library Lycopersicon esculentum genomic
            clone SL_MboI0086A08 5, genomic survey sequence.
ACCESSION CK2983576
VERSION CK2983576.1 GI:72335221
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 859)
Mueller, J.A., Buels, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,
Van Eck, J. and Stack, S.
BAC end sequencing from three Solanum lycopersicon libraries
Unpublished (2005)
Other GSSs: 197309
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@sgn.cornell.edu
Plate: 86 row: A column: 8
Seq primer: T7
Class: BAC ends
High quality sequence start: 59
High quality sequence stop: 316.
Location/Qualifiers
1..859
    /organism="Lycopersicon esculentum"
    /mol_type="genomic DNA"
    /cultivar="Heinz 1706"
    /db_xref="taxon:4081"
    /clones="SL_MboI0086A08"
    /lab_host="E. coli"
    /clone_lib="Tomato MboI BAC Library"
    /note="Vector: pBelOBAC11; Site 1: MboI"

ORIGIN
Query Match      52.9%; Score 22.2; DB 13; Length 859;
Best Local Similarity 77.1%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 5 AGCCTTGCCGCGCCCTTACAAAGACTATGAAGTA 39
    ||||| ||||| ||||| ||||| |||||
Db 354 AGGCATGCCCAACCCCATCCAAAGATTCTGAAGTA 388

Query Match      52.9%; Score 22.2; DB 13; Length 859;
Best Local Similarity 77.1%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 5 AGCCTTGCCGCGCCCTTACAAAGACTATGAAGTA 39
    ||||| ||||| ||||| ||||| |||||
Db 354 AGGCATGCCCAACCCCATCCAAAGATTCTGAAGTA 388

RESULT 82
LOCUS CK0754450 114 bp mRNA linear EST 02-AUG-2004
DEFINITION Mdfirt3050d02.y1 Mdfirt Malus x domestica cDNA clone Mdfirt3050d02 5',
            mRNA sequence.
ACCESSION CK0754450
VERSION CK0754450.1 GI:50889701
KEYWORDS EST.
SOURCE Malus x domestica
ORGANISM Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids 1; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 114)
Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A.,
Aldwinckle, H., Malnoy, M., Carroll, N., Goldsbrough, P., Orvis, K.,
Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T.,
Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I.,
Tsagarishvili, R., Kennedy, S., Waterston, R. and Wilson, R.

```


TITLE JOURNAL COMMENT

Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)
 Contact: Schuyler S. Korban
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library materials provided by: Schuyler S. Korban Library
 constructed by: K. Gasic Library sequenced by: Washington
 University Genome Sequencing Center
 WashU EST name: aaj68b01.y1
 Seq primer: -40UP from Gibco
 High quality sequence stop: 114.

FEATURES source

1. .114
 /organism="Malus x domestica"
 /mol_type="mRNA"
 /cultivar="GoldRush"
 /db_xref="taxon:3750"
 /clone="Mdft3050402"
 /tissue_type="Fruit"
 /lab_host="DH10B ampicillin resistant"
 /clone_lib="Mdft"
 /note="Vector: pBluescript II SK (+); Site 1: NotI;
 Site 2: EcoRI; Total RNA was extracted separately from
 each stage [young fruitlet (<1cm), young fruitlet (1 cm
 dia.), young fruitlet (12cm dia.), maturing fruit I,
 maturing fruit II, mature fruit], using the 'pine tree'
 method. Poly(A)+mRNA was isolated twice from total RNA
 from each stage using the Oligotex Direct mRNA kit
 (Qiagen). mRNA was reverse transcribed into double
 stranded cDNA using a modified oligo19(dT) primer with an
 identifying tag sequence (see table below). cDNA's from
 different stages were pooled in equal amounts before
 adaptor ligation. Tag identification when sequencing from
 5' end: Stage 1 (young fruitlet) insert 18(A)TCGTG; Stage
 2 (young fruitlet 1cm dia) insert 18(A)TCGTG; Stage 3
 (young fruitlet 12cm dia) insert 18(A)TCGTG; Stage 4
 (maturing fruitlet 1cm dia) insert 18(A)TCGTG; Stage 5
 (maturing fruitlet 12cm dia) insert 18(A)TCGTG; Stage 6
 (mature fruit) insert 18(A)TCGGA; Stage 7 (mature fruit) insert
 18(A)TCGTG; Tag identification when sequencing from 3',
 end: Stage 1 (young fruitlet) CACGA18(T) insert; Stage 2
 (young fruitlet 1cm dia) CACGA18(T) insert; Stage 3 (young
 fruitlet 12cm dia) ACGA18(T) insert; Stage 4 (maturing
 fruit I) TCGA18(T) insert; Stage 5 (maturing fruit II)
 TCGA18(T) insert; Stage 6 (mature fruit) ACGA18(T)
 insert. Double stranded cDNAs were size selected (more
 than 450 bp), adaptor with EcoRI adapters at both ends
 and then digested with NotI. The cDNAs were then
 directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid
 vector (Stratagene). Identification of adaptors and tags in 5'-end sequenced clones:
 (Vector) . . . TAAGTTC(End Vector) (Start
 EcoRI adaptor)GATATCGAATTCATTGTTGGG (End
 EcoRI adaptor) (Start Insert).
 Insert) (Start Tag)TCGGA(End Tag) (Start
 NotI site)Vector/CGGCCGCCACCGG. . . The total number of
 white colony forming units (cfu) in the primary library
 before amplification was 2.1x10⁶ cfu (colony forming
 units). The background of empty clones was less than 1%.
 Inserts ranged from 0.5kb to 4 kb, as determined by PCR.
 Purified plasmid DNA from the primary library was
 converted to single-stranded circles and used as a
 template for PCR amplification using the T7 and T3 priming
 sites flanking the cloned cDNA inserts. The purified PCR
 products, representing the entire cloned cDNA population,
 were used as a driver for normalization. Hybridization
 between the single-stranded library and the PCR products
 was carried out for 44 hours at 30C. Unhybridized
 single-stranded DNA circles were separated from hybridized
 DNA rendered partially double-stranded and electroporated

ORIGIN

Query Match 52.4%; Score 22; DB 8; Length 114;
 Best Local Similarity 73.7%; Pred. No. 3.1e+02;
 Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 3 CCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAA 40
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 72 CCACCATGGTGGCGTCCGTCGCTACAAAGACAGTGAAGAAA 109
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 83

AV238654 266 bp mRNA linear EST 15-NOV-2001
 AV238654 RIKEN full-length enriched, 10 day neonate skin Mus
 musculus cDNA clone 4732428J08 3', mRNA sequence.

ACCESSION

AV238654

VERSION

AV238654.1 GI:6191164

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 266)

AUTHORS

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
 Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
 Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
 Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kuakabe,M.,
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
 Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Y.,
 Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y.,
 Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tomimaga,N.,
 Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A.,
 Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE

RIKEN Mouse ESTs (Konno,H., et al. 1999)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
 Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
 Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and
 Hayashizaki,Y.
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh,M., Ktsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
 ,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)

FEATURES

Location/Qualifiers

1. .266

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="4732428J08"

/sex="mixed"

/tissue_type="skin"

/dev_stage="10 days neonate"

source

Location/Qualifiers

1. .266

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="4732428J08"

/sex="mixed"

/tissue_type="skin"

/dev_stage="10 days neonate"

source

Location/Qualifiers

1. .266

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="4732428J08"

/sex="mixed"

/tissue_type="skin"

/dev_stage="10 days neonate"

source

Location/Qualifiers

1. .266

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="4732428J08"

/sex="mixed"

/tissue_type="skin"

/dev_stage="10 days neonate"

into DH10B cells to generate the normalized library. The
 total number of clones with insert was 5.6x10⁶ cfu.
 Background of empty clones was less than 1%.

```

/lab host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 day neonate
skin"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5', GAGAGAGATTCGAGTGAATTAATTAATCCCTCCCTCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I"
```

ORIGIN

```

Query Match      52.4%; Score 22; DB 1; Length 266;
Best Local Similarity 73.7%; Pred. No. 3.7e+02;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 41
    ||||| ||||| ||||| ||||| |||||
Db 44 CCGGATTGGCAAGCTGTTCACAACTACTATTAAATAAA 81
```

RESULT 84

```

CO755512
LOCUS      Mdfrt3049p06.y1 Mdfrt Malus x domestica cDNA clone Mdfrt3049p06 5'
DEFINITION similar to TR:Q9SQZ1 Q9SQZ1 PUTATIVE 40S RIBOSOMAL PROTEIN S17. ;,
RNA sequence.
CO755512
CO755512.1 GI:50890761
```

ACCESSION

```

VERSION    Malus x domestica
KEYWORDS   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SOURCE     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
           rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
```

ORGANISM

```

Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 381)
Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A.,
Aldwinckle, H., Malnoy, M., Carroll, N., Goldsbrough, P., Orvis, K.,
Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T.,
Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I.,
Teagareihvili, R., Kennedy, S., Waterston, R. and Wilson, R.
Apple Functional Genomics grant - NSF 0321702
Unpublished (2004)
```

TITLE

```

JOURNAL    Contact: Schuyler S. Korban
COMMENT    Apple Functional Genomics grant - NSF 0321702
           Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
           Tel: 314 286 1800
           Fax: 314 286 1810
```

Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library
constructed by: K. Gasic Library sequenced by: Washington
University Genome Sequencing Center

WashU EST name: aaJ64h03.y1

Seq primer: -40UP from Gibco

High quality sequence stop: 284.

FEATURES

source

```

1..381
/organism="Malus x domestica"
/mol_type="mRNA"
/cultivar="GoldRush"
/db_xref="taxon:3750"
/clone="Mdfrt3049p06"
```

```

/tissue_type="Fruit"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdfrt"
```

```

/notes="Vector: pBluescript II SK (+); Site 1: NotI;
Site 2: EcoRI; Total RNA was extracted separately from
each stage [young fruitlet (<1cm), young fruitlet (1 cm
dia.), young fruitlet (12cm dia.), maturing fruit 1,
maturing fruit II, mature fruit], using the 'pine tree'
method. Poly(A)+mRNA was isolated twice from total RNA
from each stage using the Oligotex Direct mRNA kit
(Qiagen). mRNA was reverse transcribed into double
stranded cDNA using a modified oligo18(4T) primer with an
identifying tag sequence (see table below). cDNA's from
different stages were pooled in equal amounts before
adaptor ligation. Tag identification when sequencing from
5', end: Stage 1 (young fruitlet) insert 18(A)TCGTG; Stage
2 (young fruitlet 1cm dia) insert 18(A)TCGTG; Stage 3
(young fruitlet 12cm dia) insert 18(A)TCGTG; Stage 4
(maturing fruit I) insert 18(A)TCGA; Stage 5 (maturing
fruit II) insert 18(A)TCGA; Stage 6 (mature fruit) insert
18(A)TCGTG; Tag identification when sequencing from 3',
end: Stage 1 (young fruitlet) CAGCA18(T) insert; Stage 2
(young fruitlet 1cm dia) CAGCA18(T) insert; Stage 3 (young
fruitlet 12cm dia) ACCGA18(T) insert; Stage 4 (maturing
fruit I) TCGCA18(T) insert; Stage 5 (maturing fruit II)
TCGCA18(T) insert; Stage 6 (mature fruit) ACGCA18(T)
insert. Double stranded cDNAs were size selected (more
than 450 bp), adapted with EcoRI adaptors at both ends
and then digested with NotI. The cDNAs were then
directionally cloned into EcoRI-NotI digested pBS II SK(+)
phagemid vector(Stratagene). Identification of adaptors
and tags in 5'-end sequenced clones:
(Vector). . .TAGCTT(End Vector)(Start
EcoRI adaptor)GATACGAATTCATGTTGTGGG (End
EcoRI adaptor)(Start Insert) . . .AAAAAAAAAAAAAAAAAAAA
Insert) (Start Tag)TCGA(End Tag)(Start
NotI site/Vector)GCGCGCCACCGCGG . . . The total number of
white colony forming units (cfu) in the primary library
before amplification was 2.1x10^6 cfu (colony forming
units). The background of empty clones was less than 1%.
Inserts ranged from 0.5kb to 4 kb, as determined by PCR.
Purified plasmid DNA from the primary library was
converted to single-stranded circles and used as a
template for PCR amplification using the T7 and T3 priming
sites flanking the cloned cDNA inserts. The purified PCR
products, representing the entire cloned cDNA population,
were used as a driver for normalization. Hybridization
between the single-stranded library and the PCR products
was carried out for 44 hours at 30C. Unhybridized
single-stranded DNA circles were separated from hybridized
DNA rendered partially double-stranded and electroporated
into DH10B cells to generate the normalized library. The
total number of clones with insert was 5.6x10^6 cfu.
Background of empty clones was less than 1%."
```

ORIGIN

```

Query Match      52.4%; Score 22; DB 8; Length 381;
Best Local Similarity 73.7%; Pred. No. 3.9e+02;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 40
    ||||| ||||| ||||| ||||| |||||
Db 59 CCACATGGGTGCGTCCGTACAAAGACAGTGAAGAAA 96
```

RESULT 85

```

CO541025
LOCUS      Mdfrt3038j23.y1 Mdfrt Malus x domestica cDNA clone Mdfrt3038j23 5'
DEFINITION similar to TR:Q9SWF6 Q9SWF6 40S RIBOSOMAL PROTEIN S17. ;, mRNA
sequence.
CO541025
ACCESSION  CO541025.1 GI:50353074
VERSION
```


constructed by: K. Gasic Library sequenced by: Washington
University Genome Sequencing Center
WashU EST name: aa06b02.y1
Seq primer: -40UP from Gibco
High quality sequence stop: 461.
Location/Qualifiers

FEATURES

source

```

1..482
/organism="Malus x domestica"
/mol_type="mRNA"
/cultivar="GoldRush"
/db_xref="taxon:3750"
/clone="Mdat6016c03"
/tissue_type="Shoot internodes"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdat"
/note="Vector: pBluescript II SK (+); Site 1: NotI; Site 2: EcoRI; Total RNA was extracted separately from each stage [dormant shoot internodes, active shoot internodes, actively-growing shoot internodes], using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo(dT) primer with an identifying tag sequence (see table below). cDNA's from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (dormant shoot internodes) insert 18(A)TCGTG; Stage 2 (active shoot internodes) insert 18(A)TCGTG; Stage 3 (actively-growing shoot internodes) insert 18(A)TCGTG; Tag identification when sequencing from 3' end: Stage 1 (dormant shoot internodes) CACGA18(T) insert; Stage 2 (active shoot internodes) CACGA18(T) insert; Stage 3 (actively-growing shoot internodes) ACCGA18(T) insert; Double stranded cDNAs were size selected (more than 450 bp), adapted with EcoRI adaptors at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector(Stratagene). Identification of adaptors and tags in 5'-end sequenced clones: (Vector)...TAAGCTT(End Vector)(Start EcoRI adaptor)GATATCGAATTCATTGTGTGGG (End EcoRI adaptor)(Start Insert)...AAAAAAAAAAAAAAAAAAAA (End Insert) (Start Tag)TCGA(End Tag)(Start NotI site/Vector)GCGCGCCACCGGG... The total number of white colony forming units (cfu) in the primary library before amplification was 2.7x107 cfu (colony forming units). The background of empty clones was less than 2%. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 8x105 cfu. Background of empty clones was less than 1%."

```

ORIGIN

Query Match 52.4%; Score 22; DB 8; Length 482;
Best Local Similarity 73.7%; Pred. NO. 4.le+02;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 3 CCACGCTTGGCGGACCCCTTACAAGACTATGAGCTAA 40
|||||
Db 61 CCACATGGGTGGCTCGGTCAAAAGACAGTGAAGAAA 98
|||||

RESULT 89

CV084234

LOCUS
DEFINITION

CV084234 487 bp mRNA linear EST 26-AUG-2004
Mddb5027m14.y1 Mddb Malus x domestica cDNA clone Mddb5027m14.5,
similar to TR:Q9SWF6 Q9SWF6 40S RIBOSOMAL PROTEIN S17. ; mRNA
sequence.
CV084234
CV084234.1 GI:51563489

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Malus x domestica

Malus x domestica

REFERENCE

AUTHORS

Korban,S., Vodka,L., Liu,L., Gasic,K., Gonzales O., Hernandez,A., Aldwinckle,H., Malnoy,M., Carroll,N., Goldsbrough,P., Orvis,K., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,S., Ronko,I., Tsagarisvili,R., Kennedy,S., Waterston,R. and Wilson,R.

Apple Functional Genomics grant - NSF 0321702

Unpublished (2004)

Contact: Schuyler S. Korban

Apple Functional Genomics grant - NSF 0321702

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library

constructed by: K. Gasic Library sequenced by: Washington

University Genome Sequencing Center

WashU EST name: aa050q07.y1

Seq primer: -40UP from Gibco

High quality sequence stop: 487.

Location/Qualifiers

1..487

/organism="Malus x domestica"

/mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/clone="Mddb5027m14"

/tissue_type="Bud"

/lab_host="DH10B ampicillin resistant"

/clone_lib="Mddb"

/note="Vector: pBluescript II SK (+); Site 1: NotI; Site 2: EcoRI; Total RNA was extracted separately from each stage [dormant terminal bud, dormant lateral bud, active lateral bud], using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo(dT) primer with an identifying tag sequence (see table below). cDNAs from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (dormant terminal bud) insert 18(A)TCGTG; Stage 2 (dormant lateral bud) insert 18(A)TCGTG; Stage 3 (active lateral bud) insert 18(A)TCGTG; Tag identification when sequencing from 3' end: Stage 1 (dormant terminal bud) CACGA18(T) insert; Stage 2 (dormant lateral bud) ACCGA18(T) insert; Stage 3 (active lateral bud) ACCGA18(T) insert; Double stranded cDNAs were size selected (more than 450 bp), adapted with EcoRI adaptors at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector(Stratagene). Identification of adaptors and tags in 5'-end sequenced clones: (Vector)...TAAGCTT(End Vector)(Start EcoRI adaptor)GATATCGAATTCATTGTGTGGG (End EcoRI adaptor)(Start Insert)...AAAAAAAAAAAAAAAAAAAA (End Insert) (Start Tag)TCGA(End Tag)(Start NotI site/Vector)GCGCGCCACCGGG... The total number of white colony forming units (cfu) in the primary library before amplification was 4x10⁵ cfu (colony forming

FEATURES

source

Young Fruitlet < 1 cm, using the method described by Wang and Vodkin (Plant Molecular Biology Reporter 12:132-145, 1994). Poly(A)⁺mRNA was isolated from total RNA using the polyAtract mRNA Isolation system III (Promega). The library was prepared using the Invitrogen Life Technologies, Superscript Plasmid System with gateway technology for cDNA Synthesis and cloning. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. Sal I linker adapters were ligated to the blunt ended cDNA fragments followed by restriction with NotI. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT 1 vector. The ligated cDNA fragments were transformed into *E. coli* ElectroMax Dh10B host cells. Transformation efficiency: 10E+06 colony forming units. Average insert size by PCR: 992 bp "

[illegible]

RESULT	91
DB182714	
LOCUS	554 bp mRNA linear EST 10-DEC-2005
DEFINITION	DB182714 TLIVE2 Homo sapiens cDNA clone TLIVE2001790 5', mRNA sequence.

DB132714 GI:83513439
ACCESSION
VERSION
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

ORGANISM

Homo sapiens

Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE

1 (bases 1 to 554)

AUTHORS

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuricani, K., Wakaguri, H., Homma, H., and Homma, H.

Ishii, S., Sugiyama, T., Saito, K., Isohori, T., Kondo, H., Wagatsuma, M., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Takahashi-Fujii, A., Murakawa, K., Ishida, S., Iuchi, B., Nakai, K., Kikuchi, H., Nagai, K., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isozaki, T., and Sugano, S.

TITLE	JOURNAL
Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes	Genome Res. 16 (1), 55-65 (2006)

PURMED	COMMENT
16344560	Contact: Takao Isogai FLJ project (HRI Team) Wakayama Institute

2-6-7 Katsura Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

Email: ilj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,

FEATURES
Location/Qualifiers
pass sequencing: RAB.
Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one

```

source
1. .554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="traxn.9606"

```

```

/clone="TLIVE2001790"
/tissue_type="liver, tumor tissue"
/clone_id="TLIVE2",

```



```

insert; Stage 2 (active shoot internodes) CAGCA18(T)
insert; Stage 3 (actively-growing shoot internodes)
ACCGA18(T) insert; Double stranded cDNAs were size
selected (more than 450 bp), adapted with EcoRI adapters
at both ends and then digested with NotI. The cDNAs were
then directionally cloned into EcoRI-NotI digested pBS II
Sk(+) phagemid vector(Stratagene). Identification of
adaptors and tags in 5'-end sequenced clones:
(Vector)...TAGCTT(End Vector)(Start Insert)
EcoRI adaptor(GATATGAATTTCATTTGTTGGG (End
EcoRI adaptor)(Start Insert)...AAAAAAAAAAAAAA(End
Insert))(Start Tag)GCCGA(End Tag)(Start
NotI site/vector)GCCGCCACCACCCGG... The total number of
white colony forming units (cfu) in the primary library
before amplification was 2.7x107 cfu (colony forming
units). The background of empty clones was less than 2%.
Inserts ranged from 0.5kb to 4 kb, as determined by PCR.
Purified plasmid DNA from the primary library was
converted to single-stranded circles and used as a
template for PCR amplification using the T7 and T3 priming
sites flanking the cloned cDNA inserts. The purified PCR
products, representing the entire cloned cDNA population,
were used as a driver for normalization. Hybridization
between the single-stranded library and the PCR products
was carried out for 44 hours at 30C. Unhybridized
single-stranded DNA circles were separated from hybridized
DNA rendered partially double-stranded and electroporated
into DH10B cells to generate the normalized library. The
total number of clones with insert was 8x105 cfu.
Background of empty clones was less than 1%."

ORIGIN
Query Match 52.4%; Score 22; DB 8; Length 567;
Best Local Similarity 73.7%; Pred. No. 4.2e+02;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CCACGCTTGCCGACGCCCTTACAAGAAGACTATGAAGTAA 40
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CCACCATGGGTCCGCTCCGTACAAAGACAGTGAAGAAA 98

RESULT 93
AZ995992/c
LOCUS
DEFINITION
2M0282G06F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0282G06 F, genomic survey sequence.
ACCESSION AZ995992
VERSION AZ995992.1 GI:13867219
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
CONTACT: Robert B. Weiss
JOURNAL University of Utah Genome Center
COMMENT University of Utah
Rm. 308, Biomedical
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0282 row: G column: 06
Seq primer: CGTTGTAACACGACGCAGT

/note="Vector: pME18SFL3"

ORIGIN
Query Match 52.4%; Score 22; DB 9; Length 554;
Best Local Similarity 73.7%; Pred. No. 4.2e+02;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGACGCCCTTACAAAGACTATGAAGT 38
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 GCCCTTGCCGCGCTCCGCCCTTACACAGACTATGTGGT 192

RESULT 92
CV880385
LOCUS
DEFINITION
Mdst6018b05.y3 Mdst Malus x domestica cDNA clone Mdst6018b05 5'
similar to TR.Q9SJDO Q9SJDO 40S RIBOSOMAL PROTEIN S17. ; mRNA
sequence.
ACCESSION CV880385
VERSION CV880385.1 GI:55855593
KEYWORDS EST.
SOURCE Malus x domestica
ORGANISM Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
REFERENCE
AUTHORS Korban,S., Vockin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A.,
Alwinckle,H., Malnoy,M., Carroll,N., Goldebraugh,P., Orvis,K.,
Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I.,
Tsagarisvili,R., Kennedy,S., Waterston,R., and Wilson.R.
Apple Functional Genomics grant - NSF 0321702
Unpublished (2004)
JOURNAL
COMMENT Contact: Schuyler S. Korban
Apple Functional Genomics grant - NSF 0321702
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library materials provided by: Schuyler S. Korban Library
constructed by: K. Gasic Library sequenced by: Washington
University Genome Sequencing Center
WashU EST name: aao76a03.Y3
Seq primer: -40UP from Gibco
High quality sequence stop: 567.
Location/Qualifiers
1..567
/organism="Malus x domestica"
/mol_type="RNA"
/cultivar="GoldRush"
/db_xref="taxon:3750"
/clone="Mdst6018b05"
/tissue_type="Shoot internodes"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdst"
/note="Vector: pBluescript II SK (+); Site_1: NotI;
Site_2: EcoRI; Total RNA was extracted separately from
each stage [dormant shoot internodes, active shoot
internodes, actively-growing shoot internodes], using the
different stages were pooled in equal amounts before
total RNA from each stage using the Oligotex Direct mRNA
kit (Qiagen). mRNA was reverse transcribed into double
stranded cDNA using a modified oligo(dT) primer with an
identifying tag sequence (see table below). cDNA's from
different stages were pooled in equal amounts before
adaptor ligation. Tag identification when sequencing from
5' end: Stage 1 (dormant shoot internodes) insert
18(A)TCGTG; Stage 2 (active shoot internodes) insert
18(A)TCGTG; Stage 3 (actively-growing shoot internodes)
insert 18(A)TCGGT; Tag identification when sequencing from
3' end: Stage 1 (dormant shoot internodes) CAGCA18(T)
```

```

Class: plasmid ends
High quality sequence stop: 647.
Location/Qualifiers
1. .647
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0282G06"
/sex="Female"
/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb library"
/note="Vector: PWD42nv; Purified genomic DNA from M. Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 52.4%; Score 22; DB 11; Length 647;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CCAGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||
Db 351 CCAGGCGGCTACAAAGACTATGAAGTAAAG 322
|||

RESULT 94
CN916071 655 bp mRNA linear EST 07-JUN-2004
LOCUS
DEFINITION
CDNA clone AB0C002954H (ABOC) Crab apple mature fruit Malus sieboldii
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
Rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 655)
Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave,A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.
Location/Qualifiers
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

FEATURES
source
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

Class: plasmid ends
High quality sequence stop: 647.
Location/Qualifiers
1. .647
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0282G06"
/sex="Female"
/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb library"
/note="Vector: PWD42nv; Purified genomic DNA from M. Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 52.4%; Score 22; DB 11; Length 647;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CCAGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||
Db 351 CCAGGCGGCTACAAAGACTATGAAGTAAAG 322
|||

RESULT 94
CN916071 655 bp mRNA linear EST 07-JUN-2004
LOCUS
DEFINITION
CDNA clone AB0C002954H (ABOC) Crab apple mature fruit Malus sieboldii
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
Rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 655)
Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave,A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.
Location/Qualifiers
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

FEATURES
source
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

Class: plasmid ends
High quality sequence stop: 647.
Location/Qualifiers
1. .647
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0282G06"
/sex="Female"
/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb library"
/note="Vector: PWD42nv; Purified genomic DNA from M. Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 52.4%; Score 22; DB 11; Length 647;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CCAGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||
Db 351 CCAGGCGGCTACAAAGACTATGAAGTAAAG 322
|||

RESULT 94
CN916071 655 bp mRNA linear EST 07-JUN-2004
LOCUS
DEFINITION
CDNA clone AB0C002954H (ABOC) Crab apple mature fruit Malus sieboldii
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
Rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 655)
Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave,A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.
Location/Qualifiers
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

FEATURES
source
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

Class: plasmid ends
High quality sequence stop: 647.
Location/Qualifiers
1. .647
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0282G06"
/sex="Female"
/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb library"
/note="Vector: PWD42nv; Purified genomic DNA from M. Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 52.4%; Score 22; DB 11; Length 647;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CCAGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||
Db 351 CCAGGCGGCTACAAAGACTATGAAGTAAAG 322
|||

RESULT 94
CN916071 655 bp mRNA linear EST 07-JUN-2004
LOCUS
DEFINITION
CDNA clone AB0C002954H (ABOC) Crab apple mature fruit Malus sieboldii
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
Rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 655)
Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave,A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.
Location/Qualifiers
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

FEATURES
source
1. .655
/organism="Malus sieboldii"
/mol_type="mRNA"
/db_xref="taxon:106566"
/clone="ABOC002954"

Class: plasmid ends
High quality sequence stop: 647.
Location/Qualifiers
1. .647
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0282G06"
/sex="Female"
/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb library"
/note="Vector: PWD42nv; Purified genomic DNA from M. Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using
```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 703)
REFERENCE
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Mary Hendrix
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: NDAM1164 row: 1 column: 08
 High quality sequence start: 273
 High quality sequence stop: 378.
FEATURES
 Location/Qualifiers
 source
 1..703
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30923599"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH MGC 212"
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dt primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCA. Tissue was provided by Mary Hendrix."

FEATURES source

1..703
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30923599"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH MGC 212"
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dt primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 52.4%; Score 22; DB 8; Length 703;
 Best Local Similarity 73.7%; Pred. No. 4.4e+02;
 Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Oy 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGT 38
 |||||
 Db 328 GCCCTTGCTGCTGCGGCCCTACACAGACTATGTGTT 365

RESULT 97
BI561919
LOCUS BI561919 745 bp mRNA linear EST 05-SEP-2001
DEFINITION 603254822F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5297404 5', mRNA sequence.
ACCESSION BI561919
VERSION BI561919.1 GI:15449233
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 745)
REFERENCE
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM11752 row: k column: 05
 High quality sequence stop: 667.
FEATURES
 Location/Qualifiers
 source
 1..745
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5297404"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 97"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 52.4%; Score 22; DB 2; Length 745;
 Best Local Similarity 73.7%; Pred. No. 4.4e+02;
 Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Oy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
 |||||
 Db 608 AAGTTTCACAGAGCTTCAAAGACTCTGAAGAAAG 645

RESULT 98
BZ436965
LOCUS BZ436965 747 bp DNA linear GSS 13-DEC-2002
DEFINITION BONEV45TF BO_1.6_2_KB_tot Brassica oleracea genomic clone BONEV45, genomic survey sequence.
ACCESSION BZ436965
VERSION BZ436965.1 GI:26689633
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 747)
REFERENCE
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other_GSSs: BONEV45TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 source
 1..747
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"

FEATURES source

1..747
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"

```

BZ455062/1
LOCUS
DEFINITION
BZ455062
BZ455062.1
BZ455062
BZ455062.1
GSS.
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 771)
Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
Other GSSs: BONGH48TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
Location/Qualifiers
1..771
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONGH48"
/clone_lib="BO.1.6.2 KB tot"
/notes="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"
ORIGIN
Query Match 52.4%; Score 22; DB 11; Length 771;
Best Local Similarity 73.7%; Pred. No. 4.5e+02;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 3 CCAGCGTTGGCCAGCGCCCTTACAAAGACTATGAAGTAA 40
Db 522 CCATGCTTGGACGAGGCGCTCCTTGTGACTATGAGCGAA 485
Search completed: June 10, 2006, 19:34:27
Job time : 5242.8 secs

```


97	19.2	45.7	601	3	US-09-949-016-47026	Sequence 47026, A	c 170	19	45.2	1899	7	PCT-US95-05534-6	Sequence 6, Appli
98	19.2	45.7	630	3	US-08-612-973-31	Sequence 31, Appl	171	19	45.2	1998	3	US-09-280-799-195	Sequence 195, App
99	19.2	45.7	630	3	US-08-927-597-31	Sequence 31, Appl	172	19	45.2	1998	3	US-09-949-016-87	Sequence 87, Appl
100	19.2	45.7	630	3	US-08-928-757-31	Sequence 31, Appl	173	19	45.2	1998	5	US-09-543-679A-2552	Sequence 2552, Ap
101	19.2	45.7	630	5	US-07-973-025-31	Sequence 31, Appl	174	19	45.2	2001	3	US-09-949-016-2743	Sequence 2743, Ap
102	19.2	45.7	741	3	US-09-217-609A-29	Sequence 29, Appl	175	19	45.2	2001	3	US-09-949-016-2749	Sequence 2749, Ap
103	19.2	45.7	741	3	US-08-873-235B-29	Sequence 29, Appl	176	19	45.2	2006	2	US-07-757-390-10	Sequence 10, Appl
104	19.2	45.7	957	4	US-09-851-138C-226	Sequence 226, App	177	19	45.2	2006	2	US-07-757-390-17	Sequence 17, Appl
105	19.2	45.7	959	3	US-09-878-281A-49	Sequence 49, Appl	178	19	45.2	2006	2	US-08-442-282-10	Sequence 10, Appl
106	19.2	45.7	959	3	US-09-878-281A-51	Sequence 51, Appl	179	19	45.2	2006	2	US-08-442-282-17	Sequence 17, Appl
107	19.2	45.7	959	3	US-09-878-281A-53	Sequence 53, Appl	180	19	45.2	2006	2	US-08-442-281-10	Sequence 10, Appl
108	19.2	45.7	1455	3	US-09-248-796A-2077	Sequence 2077, Ap	181	19	45.2	2006	2	US-08-442-281-17	Sequence 17, Appl
109	19.2	45.7	1883	3	US-09-252-991A-8337	Sequence 8337, Ap	182	19	45.2	2006	2	US-08-939-727-17	Sequence 17, Appl
110	19.2	45.7	1833	3	US-09-252-991A-8371	Sequence 8371, Ap	183	19	45.2	2006	2	US-08-939-727-17	Sequence 17, Appl
111	19.2	45.7	1980	3	US-09-217-609A-8507	Sequence 8507, Ap	184	19	45.2	2006	5	US-09-543-679A-2554	Sequence 2554, App
112	19.2	45.7	2451	3	US-09-217-609A-21	Sequence 21, Appl	185	19	45.2	2006	3	US-09-949-016-15487	Sequence 15487, A
113	19.2	45.7	2451	3	US-08-973-235B-21	Sequence 21, Appl	186	19	45.2	2024	2	US-07-757-390-12	Sequence 12, Appl
114	19.2	45.7	3647	3	US-09-949-016-5484	Sequence 5484, Ap	187	19	45.2	2024	2	US-07-757-390-18	Sequence 18, Appl
115	19.2	45.7	3647	3	US-09-949-016-5485	Sequence 5485, Ap	188	19	45.2	2024	2	US-08-442-282-12	Sequence 12, Appl
116	19.2	45.7	3647	3	US-09-949-016-5486	Sequence 5486, Ap	189	19	45.2	2024	2	US-08-442-282-18	Sequence 18, Appl
117	19.2	45.7	3647	3	US-09-949-016-5487	Sequence 5487, Ap	190	19	45.2	2024	2	US-08-442-281-12	Sequence 12, Appl
118	19.2	45.7	4827	3	US-09-949-016-2774	Sequence 2774, Ap	191	19	45.2	2024	2	US-08-442-281-18	Sequence 18, Appl
119	19.2	45.7	4827	3	US-09-949-016-2775	Sequence 2775, Ap	192	19	45.2	2024	2	US-08-939-727-12	Sequence 12, Appl
120	19.2	45.7	4827	3	US-09-949-016-2776	Sequence 2776, Ap	193	19	45.2	2024	2	US-08-939-727-18	Sequence 18, Appl
121	19.2	45.7	4827	3	US-09-949-016-2777	Sequence 2777, Ap	194	19	45.2	2024	3	US-09-023-655-1315	Sequence 1315, Ap
122	19.2	45.7	4960	3	US-09-907-843-3	Sequence 3, Appli	195	19	45.2	2024	5	US-09-543-679A-2548	Sequence 2548, Ap
123	19.2	45.7	11261	3	US-09-949-016-14358	Sequence 14358, A	c 196	19	45.2	2052	2	US-08-443-104-5	Sequence 5, Appli
124	19.2	45.7	11261	3	US-09-949-016-14908	Sequence 14908, A	c 197	19	45.2	2052	2	US-08-238-130-6	Sequence 6, Appli
125	19.2	45.7	34049	3	US-09-949-016-17276	Sequence 17276, A	c 198	19	45.2	2052	2	US-08-442-859-5	Sequence 5, Appli
126	19.2	45.7	45432	3	US-09-949-002-719	Sequence 719, App	c 199	19	45.2	2052	2	US-08-398-489-5	Sequence 5, Appli
127	19.2	45.7	47419	3	US-09-949-016-11841	Sequence 11841, A	c 200	19	45.2	2052	2	US-08-894-772-1	Sequence 1, Appli
128	19.2	45.7	107800	3	US-09-949-016-15484	Sequence 15484, A	c 201	19	45.2	2052	2	US-09-207-844-1	Sequence 1, Appli
129	19.2	45.7	17820	3	US-09-949-016-13118	Sequence 13118, A	c 202	19	45.2	2052	3	US-09-252-509-1	Sequence 1, Appli
130	19.2	45.7	116425	3	US-09-949-016-11809	Sequence 11809, A	c 203	19	45.2	2052	7	PCT-US95-05534-5	Sequence 5, Appli
131	19.2	45.7	203475	3	US-09-949-016-14516	Sequence 14516, A	c 204	19	45.2	2297	3	US-09-853-753-1	Sequence 1, Appli
132	19.2	45.7	203475	3	US-09-949-016-14517	Sequence 14517, A	c 205	19	45.2	2397	3	US-10-104-047-659	Sequence 659, App
133	19.2	45.7	203475	3	US-09-949-016-14518	Sequence 14518, A	c 206	19	45.2	2397	3	US-09-949-016-2833	Sequence 2833, App
134	19.2	45.7	203475	3	US-09-949-016-14519	Sequence 14519, A	c 207	19	45.2	3737	2	US-08-784-208-1	Sequence 1, Appli
135	19.2	45.7	203475	3	US-09-949-016-17226	Sequence 17226, A	c 208	19	45.2	11355	5	US-09-543-679A-2555	Sequence 2555, App
136	19.2	45.7	203475	3	US-09-949-016-17227	Sequence 17227, A	c 209	19	45.2	32278	3	US-09-949-016-14575	Sequence 14575, A
137	19.2	45.7	203475	3	US-09-949-016-17228	Sequence 17228, A	c 210	19	45.2	44248	3	US-09-949-016-11829	Sequence 11829, A
138	19.2	45.7	203475	3	US-09-949-016-17229	Sequence 17229, A	c 211	19	45.2	44249	3	US-09-949-016-14485	Sequence 14485, A
139	19.2	45.7	678533	3	US-09-949-016-14577	Sequence 14577, A	c 212	19	45.2	44249	3	US-09-949-016-14491	Sequence 14491, A
140	19.2	45.7	678533	3	US-09-949-016-14578	Sequence 14578, A	c 213	19	45.2	79595	3	US-09-949-016-15318	Sequence 15318, A
141	19	45.2	576	2	US-08-086-428B-44	Sequence 44, Appl	c 214	19	45.2	102008	3	US-09-949-016-16617	Sequence 16617, A
142	19	45.2	576	2	US-08-468-570-44	Sequence 44, Appl	c 215	19	45.2	148794	3	US-09-949-016-12751	Sequence 12751, A
143	19	45.2	576	2	US-08-290-665A-44	Sequence 44, Appl	c 216	19	45.2	172677	3	US-09-949-016-13444	Sequence 13444, A
144	19	45.2	576	3	PCT-US95-10398-44	Sequence 44, Appl	c 217	18.8	44.8	239	3	US-09-437-457-17	Sequence 17, Appl
145	19	45.2	576	3	US-09-949-016-58674	Sequence 58674, A	c 218	18.8	44.8	601	3	US-09-949-016-79985	Sequence 79985, A
146	19	45.2	601	3	US-09-949-016-58675	Sequence 58675, A	c 219	18.8	44.8	601	3	US-09-949-016-121786	Sequence 121786, A
147	19	45.2	601	3	US-09-949-016-102574	Sequence 102574, A	c 220	18.8	44.8	601	3	US-09-949-016-131925	Sequence 131925, A
148	19	45.2	601	3	US-09-949-016-172198	Sequence 172198, A	c 221	18.8	44.8	647	3	US-09-270-767-1408	Sequence 1408, App
149	19	45.2	601	3	US-09-949-016-172199	Sequence 172199, A	c 222	18.8	44.8	647	3	US-09-270-767-16690	Sequence 16690, A
150	19	45.2	957	4	US-09-851-138C-224	Sequence 224, App	c 223	18.8	44.8	1251	2	US-09-211-930-2	Sequence 2, Appli
151	19	45.2	1188	2	US-07-757-390-11	Sequence 11, Appl	c 224	18.8	44.8	1251	2	US-09-340-993-2	Sequence 2, Appli
152	19	45.2	1188	2	US-08-442-282-11	Sequence 11, Appl	c 225	18.8	44.8	1251	3	US-09-468-442-2	Sequence 2, Appli
153	19	45.2	1188	2	US-08-442-281-11	Sequence 11, Appl	c 226	18.8	44.8	1353	2	US-09-211-930-8	Sequence 8, Appli
154	19	45.2	1188	2	US-08-939-727-11	Sequence 11, Appl	c 227	18.8	44.8	1353	2	US-09-340-993-8	Sequence 8, Appli
155	19	45.2	1260	2	US-07-757-390-9	Sequence 9, Appli	c 228	18.8	44.8	1353	3	US-09-468-442-8	Sequence 8, Appli
156	19	45.2	1260	2	US-08-442-282-9	Sequence 9, Appli	c 229	18.8	44.8	1353	3	US-09-468-442-8	Sequence 13, Appl
157	19	45.2	1260	2	US-08-442-281-9	Sequence 9, Appli	c 230	18.8	44.8	1542	3	US-09-345-473B-13	Sequence 13, Appl
158	19	45.2	1260	2	US-08-939-727-9	Sequence 9, Appli	c 231	18.8	44.8	1737	3	US-09-862-027-13	Sequence 13, Appl
159	19	45.2	1260	2	US-08-939-727-9	Sequence 9, Appli	c 232	18.8	44.8	3201	2	US-09-211-930-1	Sequence 1, Appli
160	19	45.2	1327	5	US-09-543-679A-2550	Sequence 2550, Ap	c 233	18.8	44.8	3201	3	US-09-340-993-1	Sequence 1, Appli
161	19	45.2	1351	2	US-07-421-822-1	Sequence 1, Appli	c 234	18.8	44.8	3201	3	US-09-468-442-1	Sequence 1, Appli
162	19	45.2	1351	2	US-08-421-823-1	Sequence 1, Appli	c 235	18.8	44.8	3268	3	US-09-688-188B-1	Sequence 1, Appli
163	19	45.2	1351	2	US-08-421-823-1	Sequence 1, Appli	c 236	18.8	44.8	3268	3	US-09-291-417D-1	Sequence 1, Appli
164	19	45.2	1405	5	US-09-543-679A-2553	Sequence 2553, Ap	c 237	18.8	44.8	4130	3	US-09-919-039-212	Sequence 212, App
165	19	45.2	1664	3	US-09-016-434-1220	Sequence 1220, Ap	c 238	18.8	44.8	65966	3	US-09-949-016-17152	Sequence 17152, App
166	19	45.2	1664	3	US-09-543-679A-2551	Sequence 2551, Ap	c 239	18.8	44.8	150409	3	US-09-949-016-12290	Sequence 12290, A
167	19	45.2	1899	2	US-08-443-104-6	Sequence 6, Appli	c 240	18.8	44.8	150409	3	US-09-949-016-12938	Sequence 12938, A
168	19	45.2	1899	2	US-08-442-859-6	Sequence 6, Appli	c 241	18.8	44.8	221958	3	US-09-949-016-12173	Sequence 12173, A
169	19	45.2	1899	2	US-08-398-489-6	Sequence 6, Appli	c 242	18.8	44.8	221966	3	US-09-949-016-15498	Sequence 15498, A

c 243	18.8	44.8	222691	3	US-09-949-016-11762	Sequence 11762, A	316	18.2	43.3	1275	3	US-09-328-925-80	Sequence 80, Appl
c 244	18.8	44.8	222697	3	US-09-949-016-15842	Sequence 15842, A	317	18.2	43.3	1276	3	US-09-328-925-81	Sequence 81, Appl
c 245	18.8	44.8	235064	3	US-09-949-016-15390	Sequence 13390, A	318	18.2	43.3	1427	3	US-09-577-657A-2	Sequence 2, Appl
246	18.8	44.8	312470	3	US-09-949-016-14043	Sequence 14043, A	319	18.2	43.3	1427	3	US-09-577-657A-3	Sequence 3, Appl
247	18.8	44.8	336024	3	US-09-949-016-12373	Sequence 12373, A	c 320	18.2	43.3	1472	3	US-10-101-464A-966	Sequence 966, App
248	18.8	44.8	421118	3	US-09-949-016-16297	Sequence 16297, A	321	18.2	43.3	2340	3	US-08-742-877-3	Sequence 3, Appl
c 249	18.6	44.3	354	3	US-09-513-999C-13634	Sequence 13634, A	322	18.2	43.3	2775	3	US-09-053-871A-22	Sequence 22, Appl
250	18.6	44.3	579	3	US-09-878-281A-179	Sequence 179, App	323	18.2	43.3	2792	3	US-09-118-748-1	Sequence 1, Appl
c 251	18.6	44.3	601	3	US-09-949-016-32925	Sequence 32925, App	324	18.2	43.3	2802	3	US-08-742-877-1	Sequence 1, Appl
c 252	18.6	44.3	601	3	US-09-949-016-32926	Sequence 32926, App	325	18.2	43.3	2804	3	US-08-742-877-1	Sequence 2, Appl
c 253	18.6	44.3	601	3	US-09-949-016-32927	Sequence 32927, A	326	18.2	43.3	2833	3	US-09-949-016-5252	Sequence 5252, Ap
c 254	18.6	44.3	601	3	US-09-949-016-185573	Sequence 185573, A	327	18.2	43.3	7260	3	US-09-917-254-31	Sequence 31, Appl
c 255	18.6	44.3	601	3	US-09-949-016-185574	Sequence 185574, A	328	18.2	43.3	7260	3	US-09-919-497-24	Sequence 24, Appl
c 256	18.6	44.3	601	3	US-09-949-016-185575	Sequence 185575, A	329	18.2	43.3	7260	4	US-09-880-107-3739	Sequence 3739, Ap
c 257	18.6	44.3	601	3	US-09-949-016-201189	Sequence 201189, A	330	18.2	43.3	12570	3	US-09-949-002-850	Sequence 850, App
258	18.6	44.3	603	5	US-09-974-300-798	Sequence 798, App	331	18.2	43.3	20407	3	US-09-949-002-866	Sequence 66, App
c 259	18.6	44.3	1500	3	US-09-555-704-1	Sequence 1, Appl	332	18.2	43.3	24333	5	US-09-984-429-351	Sequence 351, App
c 260	18.6	44.3	54986	3	US-09-949-016-16716	Sequence 16716, A	333	18.2	43.3	24333	5	US-09-984-429-351	Sequence 460, App
261	18.6	44.3	68035	3	US-09-949-016-16219	Sequence 16219, A	c 334	18.2	43.3	27465	3	US-09-949-016-16561	Sequence 16561, A
262	18.6	44.3	87594	3	US-09-949-016-12135	Sequence 12135, A	335	18.2	43.3	36755	3	US-09-949-016-16994	Sequence 16994, A
263	18.6	44.3	87611	3	US-09-949-016-16139	Sequence 16139, A	336	18.2	43.3	38059	3	US-09-328-925-4	Sequence 4, Appl
264	18.6	44.3	96583	4	US-09-531-120-203	Sequence 203, App	337	18.2	43.3	38059	4	US-09-880-107-2125	Sequence 2125, Ap
c 265	18.6	44.3	103987	3	US-09-949-016-12513	Sequence 12513, A	c 338	18.2	43.3	67002	3	US-09-949-016-16803	Sequence 16803, A
c 266	18.6	44.3	103988	3	US-09-949-016-17050	Sequence 17050, A	c 339	18.2	43.3	71815	3	US-09-949-016-12501	Sequence 12501, A
c 267	18.6	44.3	111677	3	US-09-949-016-16946	Sequence 16946, A	c 340	18.2	43.3	119930	3	US-09-949-016-12677	Sequence 12677, A
c 268	18.6	44.3	198632	3	US-09-949-016-12781	Sequence 12781, A	c 341	18.2	43.3	119931	3	US-09-949-016-16319	Sequence 16319, A
c 269	18.6	44.3	198637	3	US-09-949-016-17393	Sequence 17393, A	c 342	18.2	43.3	12536	3	US-09-949-016-14186	Sequence 14186, A
c 270	18.4	43.8	300	4	US-09-237-648-2393	Sequence 2393, Ap	343	18.2	43.3	136058	3	US-09-949-016-12565	Sequence 12565, A
271	18.4	43.8	576	2	US-08-086-428B-48	Sequence 48, Appl	344	18.2	43.3	136480	3	US-09-949-016-17064	Sequence 17064, A
272	18.4	43.8	576	2	US-08-468-570-48	Sequence 48, Appl	c 345	18.2	43.3	141560	3	US-09-949-016-16476	Sequence 16476, A
273	18.4	43.8	576	2	US-08-290-665A-48	Sequence 48, Appl	c 346	18.2	43.3	161652	4	US-09-497-855A-40	Sequence 40, Appl
274	18.4	43.8	576	3	US-08-466-601A-48	Sequence 48, Appl	347	18.2	43.3	611587	4	US-09-531-120-209	Sequence 209, App
275	18.4	43.8	576	7	PCT-US95-10398-48	Sequence 48, Appl	348	18.2	43.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
276	18.4	43.8	601	3	US-09-949-016-54878	Sequence 54878, A	c 349	18.2	43.3	4411529	3	US-09-103-840A-2	Sequence 1, Appl
277	18.4	43.8	721	3	US-09-533-559-2071	Sequence 2071, Ap	c 350	18.2	42.9	142	3	US-08-956-171E-4735	Sequence 4735, Ap
c 278	18.4	43.8	830	4	US-09-237-648-3596	Sequence 3596, Ap	c 351	18.2	42.9	142	3	US-08-781-986A-4735	Sequence 4735, Ap
c 279	18.4	43.8	1629	3	US-09-967-552A-27	Sequence 27, Appl	c 352	18.2	42.9	400	3	US-08-956-171E-3969	Sequence 3969, Ap
c 280	18.4	43.8	1749	3	US-09-967-552A-5	Sequence 5, Appl	c 353	18.2	42.9	400	3	US-08-781-986A-3969	Sequence 3969, Ap
c 281	18.4	43.8	1794	3	US-09-967-552A-3	Sequence 3, Appl	c 354	18.2	42.9	601	3	US-09-949-016-19854	Sequence 19854, A
c 282	18.4	43.8	2025	3	US-09-540-236-1626	Sequence 1626, Ap	c 355	18.2	42.9	601	3	US-09-949-016-118117	Sequence 118117, A
c 283	18.4	43.8	2446	3	US-09-731-166-9	Sequence 9, Appl	c 356	18.2	42.9	636	3	US-08-956-171E-3664	Sequence 3664, Ap
c 284	18.4	43.8	6336	4	US-09-880-107-1537	Sequence 1537, Ap	c 357	18.2	42.9	636	3	US-08-781-986A-3664	Sequence 3664, Ap
285	18.4	43.8	6373	3	US-08-956-171E-227	Sequence 227, App	c 358	18.2	42.9	695	4	US-09-605-703B-1873	Sequence 1873, Ap
286	18.4	43.8	6373	3	US-08-781-986A-227	Sequence 227, App	c 359	18.2	42.9	949	3	US-09-270-767-15170	Sequence 15170, A
c 287	18.4	43.8	7525	3	US-09-967-552A-72	Sequence 72, Appl	c 360	18.2	42.9	1752	3	US-09-270-767-12078	Sequence 12078, A
c 288	18.4	43.8	24553	3	US-09-949-016-16901	Sequence 16901, A	c 361	18.2	42.9	1752	3	US-09-248-796A-6141	Sequence 6141, Ap
c 289	18.4	43.8	30271	3	US-09-949-016-12796	Sequence 12796, A	c 362	18.2	42.9	2725	3	US-10-054-678-1	Sequence 1, Appl
c 290	18.4	43.8	30272	3	US-09-949-016-14006	Sequence 14006, A	c 363	18.2	42.9	2833	3	US-09-270-767-15279	Sequence 15279, A
291	18.4	43.8	47115	3	US-09-949-016-12278	Sequence 12278, A	c 364	18.2	42.9	6450	3	US-09-949-016-16596	Sequence 16596, A
292	18.4	43.8	47122	3	US-09-949-016-16520	Sequence 16520, A	c 365	18.2	42.9	7430	3	US-08-956-171E-260	Sequence 260, App
293	18.4	43.8	83391	4	US-09-531-120-189	Sequence 189, App	c 366	18.2	42.9	7430	3	US-08-781-986A-260	Sequence 260, App
294	18.4	43.8	90336	4	US-09-531-120-195	Sequence 195, App	c 367	18.2	42.9	13086	3	US-08-956-171E-16	Sequence 16, Appl
c 295	18.4	43.8	95621	3	US-09-949-016-13237	Sequence 13237, A	c 368	18.2	42.9	13086	3	US-08-781-986A-16	Sequence 16, Appl
c 296	18.4	43.8	99629	3	US-09-596-002-37	Sequence 37, Appl	c 369	18.2	42.9	19181	3	US-09-949-016-15016	Sequence 15016, A
297	18.4	43.8	109250	3	US-09-949-016-12530	Sequence 12530, A	c 370	18.2	42.9	22206	3	US-09-949-016-13901	Sequence 13901, A
298	18.4	43.8	109251	3	US-09-949-016-17321	Sequence 17321, A	c 371	18.2	42.9	23155	3	US-09-949-016-12366	Sequence 12366, A
c 299	18.4	43.8	139049	3	US-09-949-016-17030	Sequence 17030, A	c 372	18.2	42.9	26760	3	US-09-949-016-15894	Sequence 15894, A
300	18.4	43.8	139257	3	US-09-920-671-11	Sequence 11, Appl	c 373	18.2	42.9	36090	3	US-09-949-016-17424	Sequence 17424, A
c 301	18.4	43.8	161124	3	US-09-949-016-11760	Sequence 11760, A	c 374	18.2	42.9	36123	3	US-09-949-016-12402	Sequence 12402, A
c 302	18.4	43.8	228851	3	US-09-949-016-13781	Sequence 13781, A	c 375	18.2	42.9	51629	3	US-09-949-016-12483	Sequence 12483, A
c 303	18.4	43.8	392000	3	US-10-027-983-11	Sequence 11, Appl	c 376	18.2	42.9	51631	3	US-09-949-016-15496	Sequence 15496, A
304	18.2	43.3	122	3	US-09-533-559-3329	Sequence 3329, Ap	c 377	18.2	42.9	66988	3	US-09-949-016-11942	Sequence 11942, A
305	18.2	43.3	375	3	US-09-252-991A-7801	Sequence 7801, Ap	c 378	18.2	42.9	66989	3	US-09-949-016-16063	Sequence 16063, A
c 306	18.2	43.3	601	3	US-09-949-016-170595	Sequence 170595, A	c 379	18.2	42.9	75378	3	US-09-949-016-17140	Sequence 17140, A
307	18.2	43.3	601	3	US-09-949-016-186297	Sequence 186297, A	c 380	18.2	42.9	86936	3	US-09-949-016-17314	Sequence 17314, A
c 308	18.2	43.3	601	3	US-09-949-016-201190	Sequence 201190, A	c 381	18.2	42.9	94133	3	US-09-949-016-11901	Sequence 11901, A
309	18.2	43.3	1272	3	US-09-328-925-82	Sequence 82, Appl	c 382	18.2	42.9	94133	3	US-09-949-016-12713	Sequence 12713, A
310	18.2	43.3	1272	3	US-09-328-925-83	Sequence 83, Appl	c 383	18.2	42.9	94135	3	US-09-949-016-15934	Sequence 15934, A
311	18.2	43.3	1273	3	US-09-328-925-83	Sequence 83, Appl	c 384	18.2	42.9	94135	3	US-09-949-016-15935	Sequence 15935, A
312	18.2	43.3	1273	3	US-09-328-925-76	Sequence 76, Appl	c 385	18.2	42.9	94135	3	US-09-949-016-15936	Sequence 15936, A
313	18.2	43.3	1273	3	US-09-328-925-77	Sequence 77, Appl	c 386	18.2	42.9	94135	3	US-09-949-016-15937	Sequence 15937, A
314	18.2	43.3	1273	3	US-09-328-925-78	Sequence 78, Appl	c 387	18.2	42.9	94755	3	US-09-949-016-11939	Sequence 11839, A
315	18.2	43.3	1273	3	US-09-328-925-79	Sequence 79, Appl	c 388	18.2	42.9	129415	3	US-09-949-016-16997	Sequence 16997, A

535	17.6	41.9	153642	3	US-09-949-016-12174	Sequence 12174, A	c 608	17.4	41.4	644	5	US-09-989-2933A-282	Sequence 282, App
536	17.6	41.9	153643	3	US-09-949-016-15635	Sequence 15635, A	c 609	17.4	41.4	650	3	US-08-998-416-1087	Sequence 1087, App
c 537	17.6	41.9	172677	3	US-09-949-016-13444	Sequence 13444, A	c 610	17.4	41.4	701	3	US-09-735-271-266	Sequence 266, App
538	17.6	41.9	173787	3	US-09-949-016-12542	Sequence 12542, A	c 611	17.4	41.4	700	3	US-09-735-271-1254	Sequence 1254, App
539	17.6	41.9	173791	3	US-09-949-016-17302	Sequence 17302, A	c 612	17.4	41.4	700	3	US-09-735-271-1255	Sequence 1255, App
540	17.6	41.9	198942	3	US-09-949-016-13209	Sequence 13209, A	c 613	17.4	41.4	742	3	US-09-702-705-1345	Sequence 1345, App
c 541	17.6	41.9	232547	3	US-09-949-016-16603	Sequence 16603, A	c 614	17.4	41.4	742	3	US-09-736-457-1345	Sequence 1345, App
542	17.6	41.9	239527	3	US-09-949-016-15980	Sequence 15980, A	c 615	17.4	41.4	742	3	US-09-614-124B-1345	Sequence 1345, App
c 543	17.6	41.9	278866	3	US-09-949-016-13922	Sequence 13922, A	c 616	17.4	41.4	742	3	US-09-671-325-1345	Sequence 1345, App
544	17.6	41.9	278866	3	US-09-949-016-13922	Sequence 13922, A	c 617	17.4	41.4	742	3	US-09-671-325-1345	Sequence 1345, App
c 545	17.6	41.9	278866	3	US-09-949-016-13924	Sequence 13924, A	c 618	17.4	41.4	742	3	US-09-671-325-1345	Sequence 1345, App
546	17.6	41.9	278866	3	US-09-949-016-13925	Sequence 13925, A	c 619	17.4	41.4	742	3	US-09-671-325-1345	Sequence 1345, App
c 547	17.6	41.9	278866	3	US-09-949-016-13925	Sequence 13925, A	c 620	17.4	41.4	742	3	US-09-671-325-1345	Sequence 1345, App
548	17.6	41.9	278866	3	US-09-949-016-14699	Sequence 14699, A	c 621	17.4	41.4	1452	3	US-08-887-534A-50	Sequence 50, Appl
c 549	17.6	41.9	278866	3	US-09-949-016-14699	Sequence 14699, A	c 622	17.4	41.4	1452	3	US-09-527-431-50	Sequence 50, Appl
550	17.6	41.9	278866	3	US-09-949-016-14701	Sequence 14701, A	c 623	17.4	41.4	1452	3	US-09-446-861-50	Sequence 50, Appl
c 551	17.6	41.9	278866	3	US-09-949-016-14702	Sequence 14702, A	c 624	17.4	41.4	1452	3	US-09-446-861-50	Sequence 50, Appl
552	17.6	41.9	278866	3	US-09-949-016-14703	Sequence 14703, A	c 625	17.4	41.4	1635	3	US-09-107-532A-941	Sequence 941, App
c 553	17.6	41.9	300402	3	US-09-949-016-13632	Sequence 13632, A	c 626	17.4	41.4	1635	3	US-09-446-861-50	Sequence 1439, App
554	17.6	41.9	30416	3	US-09-949-016-16923	Sequence 16923, A	c 627	17.4	41.4	1692	3	US-09-601-198-63	Sequence 63, Appl
c 555	17.6	41.9	524032	3	US-09-949-016-16928	Sequence 16928, A	c 628	17.4	41.4	2158	3	US-09-318-448-4	Sequence 4, Appl
556	17.6	41.9	524032	3	US-09-949-016-16928	Sequence 16928, A	c 629	17.4	41.4	2158	3	US-09-962-665-12	Sequence 12, Appl
c 557	17.6	41.9	524032	3	US-09-949-016-16929	Sequence 16929, A	c 630	17.4	41.4	2158	3	US-09-962-665-12	Sequence 12, Appl
558	17.6	41.9	529885	3	US-09-949-016-16931	Sequence 16931, A	c 631	17.4	41.4	2158	3	US-09-962-665-12	Sequence 12, Appl
c 559	17.6	41.9	529885	3	US-09-949-016-16931	Sequence 16931, A	c 632	17.4	41.4	2158	3	US-09-962-665-12	Sequence 12, Appl
560	17.6	41.9	529885	3	US-09-949-016-14341	Sequence 14341, A	c 633	17.4	41.4	2295	3	US-09-221-017B-736	Sequence 736, App
c 561	17.6	41.9	529885	3	US-09-949-016-14342	Sequence 14342, A	c 634	17.4	41.4	2295	3	US-09-221-017B-736	Sequence 736, App
562	17.6	41.9	529885	3	US-09-949-016-14343	Sequence 14343, A	c 635	17.4	41.4	2840	3	US-09-799-451-392	Sequence 392, App
c 563	17.6	41.9	529885	3	US-09-949-016-14344	Sequence 14344, A	c 636	17.4	41.4	3268	3	US-10-104-047-259	Sequence 259, App
564	17.6	41.9	529885	3	US-09-949-016-14345	Sequence 14345, A	c 637	17.4	41.4	4580	3	US-10-104-047-259	Sequence 259, App
c 565	17.6	41.9	529885	3	US-09-949-016-14346	Sequence 14346, A	c 638	17.4	41.4	4580	3	US-10-120-988-275	Sequence 275, App
566	17.6	41.9	529885	3	US-09-949-016-14347	Sequence 14347, A	c 639	17.4	41.4	4682	3	US-09-774-528-274	Sequence 274, App
c 567	17.6	41.9	529885	3	US-09-949-016-14347	Sequence 14347, A	c 640	17.4	41.4	4682	3	US-09-774-528-274	Sequence 274, App
568	17.6	41.9	1830121	3	US-09-557-884-1	Sequence 1, Appl	c 641	17.4	41.4	4682	3	US-10-120-988-274	Sequence 274, App
c 569	17.6	41.9	1830121	3	US-09-557-884-1	Sequence 1, Appl	c 642	17.4	41.4	4682	3	US-10-120-988-274	Sequence 274, App
570	17.6	41.9	4403765	3	US-10-158-865-1	Sequence 1, Appl	c 643	17.4	41.4	12835	3	US-09-949-016-13334	Sequence 13334, A
c 571	17.6	41.9	4411529	3	US-09-103-840A-2	Sequence 2, Appl	c 644	17.4	41.4	13518	3	US-09-949-016-16687	Sequence 16687, A
572	17.4	41.4	273	3	US-09-313-294A-3131	Sequence 1, Appl	c 645	17.4	41.4	15077	3	US-09-949-016-12441	Sequence 12441, A
c 573	17.4	41.4	286	3	US-09-270-767-3318	Sequence 3131, App	c 646	17.4	41.4	15077	3	US-09-949-016-12441	Sequence 12441, A
574	17.4	41.4	286	3	US-09-270-767-18600	Sequence 18600, A	c 647	17.4	41.4	15131	3	US-09-949-016-16805	Sequence 16805, A
c 575	17.4	41.4	494	3	US-09-621-376-2866	Sequence 2866, App	c 648	17.4	41.4	16052	3	US-09-949-016-13181	Sequence 13181, A
576	17.4	41.4	566	3	US-09-533-559-4664	Sequence 4664, App	c 649	17.4	41.4	23796	3	US-09-949-016-17581	Sequence 17581, A
c 577	17.4	41.4	601	3	US-09-949-016-25936	Sequence 25936, A	c 650	17.4	41.4	29293	3	US-09-949-016-17581	Sequence 17581, A
578	17.4	41.4	601	3	US-09-949-016-54133	Sequence 54133, A	c 651	17.4	41.4	29293	3	US-09-949-016-16363	Sequence 16363, A
c 579	17.4	41.4	601	3	US-09-949-016-58587	Sequence 58587, A	c 652	17.4	41.4	30246	3	US-09-949-016-13442	Sequence 13442, A
580	17.4	41.4	601	3	US-09-949-016-60272	Sequence 60272, A	c 653	17.4	41.4	30246	3	US-08-956-171B-56	Sequence 56, Appl
c 581	17.4	41.4	601	3	US-09-949-016-60273	Sequence 60273, A	c 654	17.4	41.4	31422	3	US-08-781-986A-56	Sequence 56, Appl
582	17.4	41.4	601	3	US-09-949-016-74538	Sequence 74538, A	c 655	17.4	41.4	31422	3	US-09-914-286-2	Sequence 2, Appl
c 583	17.4	41.4	601	3	US-09-949-016-77623	Sequence 77623, A	c 656	17.4	41.4	37068	3	US-09-949-016-12543	Sequence 12543, A
584	17.4	41.4	601	3	US-09-949-016-80254	Sequence 80254, A	c 657	17.4	41.4	37068	3	US-09-949-016-17376	Sequence 17376, A
c 585	17.4	41.4	601	3	US-09-949-016-82627	Sequence 82627, A	c 658	17.4	41.4	37304	3	US-09-949-016-11862	Sequence 11862, A
586	17.4	41.4	601	3	US-09-949-016-85515	Sequence 85515, A	c 659	17.4	41.4	43726	3	US-09-949-016-17578	Sequence 17578, A
c 587	17.4	41.4	601	3	US-09-949-016-85516	Sequence 85516, A	c 660	17.4	41.4	47110	3	US-09-949-016-17609	Sequence 17609, A
588	17.4	41.4	601	3	US-09-949-016-173650	Sequence 173650, A	c 661	17.4	41.4	48316	3	US-09-949-016-12853	Sequence 12853, A
c 589	17.4	41.4	601	3	US-09-949-016-192367	Sequence 192367, A	c 662	17.4	41.4	48316	3	US-09-949-016-13405	Sequence 13405, A
590	17.4	41.4	601	3	US-09-949-016-195434	Sequence 195434, A	c 663	17.4	41.4	50950	3	US-09-949-016-16659	Sequence 16659, A
c 591	17.4	41.4	601	3	US-09-949-016-195435	Sequence 195435, A	c 664	17.4	41.4	63385	3	US-09-949-016-14856	Sequence 14856, A
592	17.4	41.4	601	3	US-09-949-016-195436	Sequence 195436, A	c 665	17.4	41.4	72742	3	US-09-949-016-15365	Sequence 15365, A
c 593	17.4	41.4	601	3	US-09-949-016-196159	Sequence 196159, A	c 666	17.4	41.4	92334	3	US-09-949-016-16361	Sequence 16361, A
594	17.4	41.4	601	3	US-09-949-016-196160	Sequence 196160, A	c 667	17.4	41.4	92363	3	US-09-949-016-13920	Sequence 13920, A
c 595	17.4	41.4	601	3	US-09-949-016-201338	Sequence 201338, A	c 668	17.4	41.4	95223	4	US-09-531-120-188	Sequence 188, App
596	17.4	41.4	601	3	US-09-949-016-201452	Sequence 201452, A	c 669	17.4	41.4	99916	3	US-09-816-095-3	Sequence 3, Appl
c 597	17.4	41.4	601	3	US-09-949-016-206935	Sequence 206935, A	c 670	17.4	41.4	112114	3	US-09-949-016-17292	Sequence 17292, A
598	17.4	41.4	644	3	US-09-991-181-282	Sequence 282, App	c 671	17.4	41.4	120213	3	US-09-949-016-13304	Sequence 13304, A
c 599	17.4	41.4	644	3	US-09-990-444-282	Sequence 282, App	c 672	17.4	41.4	120217	3	US-09-949-016-12260	Sequence 12260, A
600	17.4	41.4	644	3	US-09-997-333-282	Sequence 282, App	c 673	17.4	41.4	121068	3	US-09-949-016-14138	Sequence 14138, A
c 601	17.4	41.4	644	3	US-09-992-598-282	Sequence 282, App	c 674	17.4	41.4	128175	3	US-09-949-016-16268	Sequence 16268, A
602	17.4	41.4	644	4	US-09-989-735-282	Sequence 282, App	c 675	17.4	41.4	128516	3	US-09-949-016-13501	Sequence 13501, A
c 603	17.4	41.4	644	5	US-09-989-726-282	Sequence 282, App	c 676	17.4	41.4	133157	3	US-09-949-016-12541	Sequence 12541, A
604	17.4	41.4	644	5	US-09-997-514-282	Sequence 282, App	c 677	17.4	41.4	145241	3	US-09-949-016-17394	Sequence 17394, A
c 605	17.4	41.4	644	5	US-09-989-728-282	Sequence 282, App	c 678	17.4	41.4	145241	3	US-09-949-016-17395	Sequence 17395, A
606	17.4	41.4	644	5	US-09-997-349-282	Sequence 282, App	c 679	17.4	41.4	154605	3	US-09-949-016-11894	Sequence 11894, A
c 607	17.4	41.4	644	5	US-09-997-653-282	Sequence 282, App	c 680	17.4	41.4	163317	4	US-09-531-120-212	Sequence 212, App

c 681	17.4	41.4	173787	3	US-09-949-016-12542	Sequence 12542, A	754	17.2	41.0	4078	3	US-09-016-434-870	Sequence 870, App
c 682	17.4	41.4	173791	3	US-09-949-016-17302	Sequence 17302, A	755	17.2	41.0	4931	3	US-08-726-320-2	Sequence 2, Appli
c 683	17.4	41.4	187136	3	US-09-949-016-17231	Sequence 17231, A	756	17.2	41.0	4931	3	US-09-208-716-2	Sequence 2, Appli
c 684	17.4	41.4	187136	3	US-09-949-016-17231	Sequence 17231, A	757	17.2	41.0	5391	3	US-09-415-161-2	Sequence 2, Appli
c 685	17.4	41.4	197131	3	US-09-949-016-16170	Sequence 12675, A	758	17.2	41.0	5391	3	US-09-415-161-16	Sequence 16, Appl
c 685	17.4	41.4	197131	3	US-09-949-016-16170	Sequence 12675, A	759	17.2	41.0	5407	3	US-09-415-161-17	Sequence 17, Appl
c 687	17.4	41.4	200663	3	US-09-949-016-12369	Sequence 121569, A	c 760	17.2	41.0	5887	2	US-08-380-403A-3	Sequence 3, Appli
c 688	17.4	41.4	254405	3	US-09-949-016-12369	Sequence 121569, A	c 761	17.2	41.0	5887	2	US-08-895-628-3	Sequence 3, Appli
c 689	17.4	41.4	283538	3	US-09-949-016-13506	Sequence 14381, A	762	17.2	41.0	5687	2	US-08-895-628-3	Sequence 3, Appli
c 690	17.4	41.4	340380	3	US-09-949-016-14179	Sequence 14179, A	763	17.2	41.0	5687	2	US-08-895-628-3	Sequence 3, Appli
c 691	17.4	41.4	1830121	3	US-09-557-884-1	Sequence 1, Appli	764	17.2	41.0	7001	3	US-09-949-016-13062	Sequence 13062, A
c 692	17.4	41.4	1830121	3	US-09-643-990A-1	Sequence 1, Appli	c 765	17.2	41.0	14507	3	US-08-785-150-1	Sequence 1, Appli
c 693	17.4	41.4	1830121	3	US-10-158-865-1	Sequence 1, Appli	c 766	17.2	41.0	14507	3	US-09-660-299-1	Sequence 1, Appli
c 694	17.2	41.0	218	3	US-08-956-171E-1038	Sequence 1038, Ap	c 767	17.2	41.0	14507	3	US-09-435-377-1	Sequence 1, Appli
c 695	17.2	41.0	218	3	US-08-781-986A-1038	Sequence 1038, Ap	c 768	17.2	41.0	14507	3	US-09-973-928-1	Sequence 1, Appli
c 696	17.2	41.0	237	3	US-09-270-767-13176	Sequence 12176, A	c 769	17.2	41.0	16207	3	US-09-949-016-16747	Sequence 16747, A
c 697	17.2	41.0	267	3	US-09-513-990-31605	Sequence 31605, A	c 770	17.2	41.0	17644	3	US-09-949-016-16342	Sequence 16342, A
c 698	17.2	41.0	601	3	US-09-949-016-41606	Sequence 41606, A	c 771	17.2	41.0	17645	3	US-09-949-016-11805	Sequence 11805, A
c 699	17.2	41.0	601	3	US-09-949-016-57190	Sequence 57190, A	c 772	17.2	41.0	17645	3	US-09-949-016-14188	Sequence 14188, A
c 700	17.2	41.0	601	3	US-09-949-016-77806	Sequence 77806, A	c 773	17.2	41.0	17846	3	US-08-956-171E-109	Sequence 109, App
c 701	17.2	41.0	601	3	US-09-949-016-79831	Sequence 79831, A	c 774	17.2	41.0	17846	3	US-08-781-986A-109	Sequence 109, App
c 702	17.2	41.0	601	3	US-09-949-016-109622	Sequence 109622, A	c 775	17.2	41.0	22786	3	US-09-949-016-13399	Sequence 13399, A
c 703	17.2	41.0	601	3	US-09-949-016-109623	Sequence 109623, A	c 776	17.2	41.0	24841	3	US-09-949-016-14009	Sequence 14009, A
c 704	17.2	41.0	601	3	US-09-949-016-121958	Sequence 121958, A	c 777	17.2	41.0	25231	3	US-09-949-016-12041	Sequence 12041, A
c 705	17.2	41.0	601	3	US-09-949-016-152760	Sequence 152760, A	c 778	17.2	41.0	29899	3	US-09-902-540-1265	Sequence 1265, Ap
c 706	17.2	41.0	601	3	US-09-949-016-152761	Sequence 152761, A	c 779	17.2	41.0	36417	3	US-09-949-016-14765	Sequence 14765, Ap
c 707	17.2	41.0	601	3	US-09-949-016-153456	Sequence 153456, A	c 780	17.2	41.0	41322	3	US-08-311-731A-123	Sequence 123, App
c 708	17.2	41.0	601	3	US-09-949-016-190959	Sequence 190959, A	c 781	17.2	41.0	42975	3	US-10-024-396-13	Sequence 13, Appl
c 709	17.2	41.0	601	3	US-09-949-016-199650	Sequence 199650, A	c 782	17.2	41.0	45845	3	US-09-949-016-11965	Sequence 11965, A
c 710	17.2	41.0	615	4	US-09-605-703B-1549	Sequence 1549, Ap	c 783	17.2	41.0	46899	2	US-09-927-091-6	Sequence 6, Appli
c 711	17.2	41.0	696	3	US-08-998-416-1117	Sequence 1117, Ap	c 784	17.2	41.0	49744	3	US-08-471-119A-1	Sequence 1, Appli
c 712	17.2	41.0	756	5	US-09-974-300-4269	Sequence 4269, Ap	c 785	17.2	41.0	49744	3	US-09-927-091-4	Sequence 4, Appli
c 713	17.2	41.0	795	3	US-09-902-540-5464	Sequence 5464, Ap	c 786	17.2	41.0	54531	3	US-09-949-016-16267	Sequence 16267, A
c 714	17.2	41.0	833	3	US-08-998-416-366	Sequence 366, App	c 787	17.2	41.0	70947	3	US-09-949-016-15165	Sequence 15165, A
c 715	17.2	41.0	852	3	US-09-248-796A-2002	Sequence 2002, Ap	c 788	17.2	41.0	71645	3	US-09-949-016-12126	Sequence 12126, A
c 716	17.2	41.0	864	3	US-09-603-208A-143	Sequence 143, App	c 789	17.2	41.0	71651	3	US-09-949-016-17258	Sequence 17258, A
c 717	17.2	41.0	990	3	US-09-134-000C-1367	Sequence 1367, Ap	c 790	17.2	41.0	75674	3	US-09-949-016-17597	Sequence 17597, A
c 718	17.2	41.0	1221	3	US-09-248-796A-4117	Sequence 4117, Ap	c 791	17.2	41.0	95318	3	US-09-949-016-11784	Sequence 11784, A
c 719	17.2	41.0	1251	3	US-09-328-352-2549	Sequence 2549, Ap	c 792	17.2	41.0	95318	3	US-09-949-016-13998	Sequence 13998, A
c 720	17.2	41.0	1308	2	US-08-404-531B-1	Sequence 1, Appli	c 793	17.2	41.0	107937	3	US-09-949-016-17192	Sequence 17192, A
c 721	17.2	41.0	1308	2	US-08-404-531B-2	Sequence 2, Appli	c 794	17.2	41.0	130563	3	US-09-949-016-12273	Sequence 12273, A
c 722	17.2	41.0	1308	3	US-08-476-900A-1	Sequence 1, Appli	c 795	17.2	41.0	130563	3	US-09-949-016-16050	Sequence 16050, A
c 723	17.2	41.0	1308	3	US-08-476-900A-2	Sequence 2, Appli	c 796	17.2	41.0	135171	3	US-09-949-016-15617	Sequence 15617, A
c 724	17.2	41.0	1308	3	US-08-488-546A-1	Sequence 1, Appli	c 797	17.2	41.0	147840	3	US-09-949-016-15236	Sequence 15236, A
c 725	17.2	41.0	1308	3	US-08-488-546A-2	Sequence 2, Appli	c 798	17.2	41.0	155019	3	US-09-949-016-16029	Sequence 16029, A
c 726	17.2	41.0	1352	3	US-09-949-016-1320	Sequence 1320, Ap	c 799	17.2	41.0	187916	3	US-09-949-016-12980	Sequence 12980, A
c 727	17.2	41.0	1399	3	US-09-380-287A-1	Sequence 1, Appli	c 800	17.2	41.0	189560	3	US-09-949-016-17202	Sequence 17202, A
c 728	17.2	41.0	1608	3	US-09-328-352-389	Sequence 389, App	c 801	17.2	41.0	194537	3	US-09-949-016-12928	Sequence 12928, A
c 729	17.2	41.0	1633	2	US-07-866-979-5	Sequence 5, Appli	c 802	17.2	41.0	201529	3	US-09-949-016-12740	Sequence 12740, A
c 730	17.2	41.0	1633	2	US-08-466-906B-5	Sequence 5, Appli	c 803	17.2	41.0	218940	3	US-09-949-016-17539	Sequence 17539, A
c 731	17.2	41.0	1633	3	US-08-706-281A-5	Sequence 5, Appli	c 804	17.2	41.0	260286	3	US-09-949-016-17037	Sequence 17037, A
c 732	17.2	41.0	1633	3	US-09-201-746-5	Sequence 5, Appli	c 805	17.2	41.0	260286	3	US-09-949-016-12106	Sequence 12106, A
c 733	17.2	41.0	1633	3	US-09-097-231-5	Sequence 5, Appli	c 806	17.2	41.0	325034	3	US-09-949-016-14957	Sequence 14957, A
c 734	17.2	41.0	1633	3	US-09-353-099-5	Sequence 5, Appli	c 807	17.2	41.0	373182	3	US-09-949-016-17371	Sequence 17371, A
c 735	17.2	41.0	1633	3	US-09-016-434-1343	Sequence 1343, Ap	c 808	17.2	41.0	373694	3	US-09-949-016-12062	Sequence 12062, A
c 736	17.2	41.0	1674	3	US-09-712-768-1	Sequence 1, Appli	c 809	17.2	41.0	389504	3	US-09-949-016-11774	Sequence 11774, A
c 737	17.2	41.0	1979	3	US-09-685-462-3	Sequence 3, Appli	c 810	17.2	41.0	678533	3	US-09-949-016-14577	Sequence 14577, A
c 738	17.2	41.0	2028	2	US-09-211-930-12	Sequence 12, Appl	c 811	17.2	41.0	678533	3	US-09-949-016-14578	Sequence 14578, A
c 739	17.2	41.0	2028	3	US-09-340-993-12	Sequence 12, Appl	c 812	17	40.5	186	3	US-09-016-434-369	Sequence 369, App
c 740	17.2	41.0	2028	2	US-09-468-442-12	Sequence 12, Appl	c 813	17	40.5	271	3	US-09-533-559-5657	Sequence 5657, Ap
c 741	17.2	41.0	2454	2	US-08-404-531B-32	Sequence 32, Appl	c 814	17	40.5	284	3	US-09-313-294A-1342	Sequence 1342, Ap
c 742	17.2	41.0	2454	3	US-08-476-900A-32	Sequence 32, Appl	c 815	17	40.5	327	3	US-09-902-540-4380	Sequence 4380, Ap
c 743	17.2	41.0	2454	3	US-08-488-546A-32	Sequence 32, Appl	c 816	17	40.5	346	5	US-09-974-300-7862	Sequence 7862, Ap
c 744	17.2	41.0	2485	3	US-09-976-594-1007	Sequence 1007, Ap	c 817	17	40.5	382	3	US-09-513-999C-30573	Sequence 30573, A
c 745	17.2	41.0	2538	3	US-09-527-084A-7	Sequence 7, Appli	c 818	17	40.5	396	3	US-08-956-171E-3074	Sequence 3074, Ap
c 746	17.2	41.0	2747	3	US-09-949-016-1657	Sequence 1657, Ap	c 819	17	40.5	396	3	US-08-781-986A-3074	Sequence 3074, Ap
c 747	17.2	41.0	2945	3	US-09-527-084A-5	Sequence 5, Appli	c 820	17	40.5	400	3	US-08-956-171E-3853	Sequence 3853, Ap
c 748	17.2	41.0	3089	3	US-09-016-434-1330	Sequence 1330, Ap	c 821	17	40.5	400	3	US-08-781-986A-3074	Sequence 3074, Ap
c 749	17.2	41.0	3089	4	US-09-880-107-3687	Sequence 3687, Ap	c 822	17	40.5	409	3	US-09-621-976-16228	Sequence 16228, A
c 750	17.2	41.0	3218	2	US-08-677-862-1	Sequence 1, Appli	c 823	17	40.5	489	5	US-09-974-300-5907	Sequence 5907, Ap
c 751	17.2	41.0	3218	2	US-09-252-571-1	Sequence 1, Appli	c 824	17	40.5	493	3	US-10-131-827-8883	Sequence 8883, Ap
c 752	17.2	41.0	3218	3	US-09-434-065-1	Sequence 1, Appli	c 825	17	40.5	601	3	US-09-949-016-49757	Sequence 49757, A
c 753	17.2	41.0	3880	3	US-09-221-017B-1071	Sequence 1071, Ap	c 826	17	40.5	601	3	US-09-949-016-55242	Sequence 55242, A

C 827	17	40.5	601	3	US-09-949-016-65408	Sequence 65408, A	C 900	17	40.5	3134	3	US-08-374-219B-7	Sequence 7, Appli
C 828	17	40.5	601	3	US-09-949-016-76595	Sequence 76595, A	C 901	17	40.5	3276	4	US-10-094-749-279	Sequence 279, App
C 829	17	40.5	601	3	US-09-949-016-77543	Sequence 77543, A	C 902	17	40.5	3646	3	US-09-710-729-4401	Sequence 4401, Ap
C 830	17	40.5	601	3	US-09-949-016-78852	Sequence 78852, A	C 903	17	40.5	3744	3	US-08-961-527-263	Sequence 263, App
C 831	17	40.5	601	3	US-09-949-016-78853	Sequence 78853, A	C 904	17	40.5	3795	3	US-10-002-344A-25	Sequence 25, Appli
C 832	17	40.5	601	3	US-09-949-016-107890	Sequence 107890, A	C 905	17	40.5	3921	2	US-08-567-375-3	Sequence 3, Appli
C 833	17	40.5	601	3	US-09-949-016-113399	Sequence 113399, A	C 906	17	40.5	3921	2	US-08-587-680A-3	Sequence 3, Appli
C 834	17	40.5	601	3	US-09-949-016-113400	Sequence 113400, A	C 907	17	40.5	3922	3	US-09-949-016-13969	Sequence 3969, Ap
C 835	17	40.5	601	3	US-09-949-016-123964	Sequence 123964, A	C 908	17	40.5	3923	3	US-09-949-016-844	Sequence 844, App
C 836	17	40.5	601	3	US-09-949-016-129176	Sequence 129176, A	C 909	17	40.5	5153	3	US-09-774-528-21	Sequence 21, Appli
C 837	17	40.5	601	3	US-09-949-016-129177	Sequence 129177, A	C 910	17	40.5	5153	3	US-10-120-988-21	Sequence 21, Appli
C 838	17	40.5	601	3	US-09-949-016-171946	Sequence 171946, A	C 911	17	40.5	5207	3	US-09-774-528-20	Sequence 20, Appli
C 839	17	40.5	601	3	US-09-949-016-180812	Sequence 180812, A	C 912	17	40.5	5207	3	US-10-120-988-20	Sequence 20, Appli
C 840	17	40.5	650	3	US-09-270-767-1808	Sequence 1808, Ap	C 913	17	40.5	5455	3	US-09-578-458-15	Sequence 15, Appli
C 841	17	40.5	650	3	US-09-270-767-17090	Sequence 17090, A	C 914	17	40.5	5455	3	US-09-362-336A-1	Sequence 1, Appli
C 842	17	40.5	669	3	US-08-956-171E-388	Sequence 388, App	C 915	17	40.5	5992	2	US-08-475-891A-3	Sequence 3, Appli
C 843	17	40.5	669	3	US-08-781-866A-388	Sequence 388, App	C 916	17	40.5	6014	5	US-09-984-429-368	Sequence 368, App
C 844	17	40.5	741	3	US-09-902-540-3629	Sequence 3629, Ap	C 917	17	40.5	6028	3	US-09-362-336A-3	Sequence 3, Appli
C 845	17	40.5	743	3	US-09-342-681C-11	Sequence 11, Appli	C 918	17	40.5	9038	3	US-09-949-016-13523	Sequence 13523, A
C 846	17	40.5	842	3	US-09-154-750A-72	Sequence 72, Appli	C 919	17	40.5	9210	3	US-09-712-363-100	Sequence 100, App
C 847	17	40.5	851	4	US-09-297-648-2987	Sequence 2987, Ap	C 920	17	40.5	9464	3	US-09-593-580B-1	Sequence 1, Appli
C 848	17	40.5	851	4	US-09-297-648-2988	Sequence 2988, Ap	C 921	17	40.5	9464	3	US-09-738-847-1	Sequence 1, Appli
C 849	17	40.5	879	3	US-09-976-594-987	Sequence 987, App	C 922	17	40.5	9464	3	US-09-949-002-658	Sequence 1, Appli
C 850	17	40.5	913	3	US-09-533-559-3441	Sequence 3441, Ap	C 923	17	40.5	11378	3	US-09-378-810-1	Sequence 1, Appli
C 851	17	40.5	999	3	US-09-252-991A-14199	Sequence 14199, A	C 924	17	40.5	15000	3	US-09-949-016-13815	Sequence 13815, A
C 852	17	40.5	1052	3	US-09-270-767-27177	Sequence 27177, A	C 925	17	40.5	15206	3	US-09-949-016-13585	Sequence 13585, A
C 853	17	40.5	1092	3	US-09-252-991A-16279	Sequence 16279, A	C 926	17	40.5	15206	3	US-09-949-016-13586	Sequence 13586, A
C 854	17	40.5	1141	3	US-09-617-174C-1	Sequence 1, Appli	C 927	17	40.5	15598	3	US-08-956-171E-82	Sequence 82, Appli
C 855	17	40.5	1179	3	US-09-491-577-81	Sequence 81, Appli	C 928	17	40.5	15598	3	US-08-781-891A-82	Sequence 82, Appli
C 856	17	40.5	1194	3	US-09-134-000C-2759	Sequence 2759, Ap	C 929	17	40.5	16442	3	US-08-781-891-208	Sequence 208, App
C 857	17	40.5	1305	3	US-09-540-236-933	Sequence 933, App	C 930	17	40.5	16442	3	US-09-618-166-208	Sequence 208, App
C 858	17	40.5	1438	3	US-09-187-331-4	Sequence 4, Appli	C 931	17	40.5	17198	3	US-09-949-016-13776	Sequence 13776, A
C 859	17	40.5	1438	3	US-09-470-946-4	Sequence 4, Appli	C 932	17	40.5	17318	3	US-09-949-016-13817	Sequence 13817, A
C 860	17	40.5	1438	3	US-08-928-069-5	Sequence 5, Appli	C 933	17	40.5	17591	3	US-09-949-002-658	Sequence 658, App
C 861	17	40.5	1438	3	US-08-828-683A-5	Sequence 5, Appli	C 934	17	40.5	17591	3	US-09-949-002-828	Sequence 828, App
C 862	17	40.5	1438	3	US-09-339-159B-3	Sequence 3, Appli	C 935	17	40.5	18082	3	US-09-949-016-12586	Sequence 12586, A
C 863	17	40.5	1470	3	US-09-339-159B-1	Sequence 1, Appli	C 936	17	40.5	18082	3	US-09-949-016-15711	Sequence 15711, A
C 864	17	40.5	1579	4	US-10-036-041-8	Sequence 8, Appli	C 937	17	40.5	18474	3	US-09-949-016-17559	Sequence 17559, A
C 865	17	40.5	1581	3	US-09-134-000C-3099	Sequence 3099, Ap	C 938	17	40.5	18848	3	US-09-902-540-1174	Sequence 1174, Ap
C 866	17	40.5	1791	3	US-09-270-767-11575	Sequence 11575, A	C 939	17	40.5	25190	3	US-09-949-016-15906	Sequence 15906, A
C 867	17	40.5	1959	3	US-09-252-991A-14026	Sequence 14026, A	C 940	17	40.5	27295	3	US-09-949-016-16605	Sequence 16605, A
C 868	17	40.5	2001	3	US-09-799-451-695	Sequence 695, App	C 941	17	40.5	29272	3	US-09-902-540-1217	Sequence 1217, Ap
C 869	17	40.5	2004	4	US-10-094-749-1462	Sequence 1462, Ap	C 942	17	40.5	29954	3	US-09-949-016-13808	Sequence 13808, A
C 870	17	40.5	2174	3	US-09-634-238-54	Sequence 54, Appli	C 943	17	40.5	31040	3	US-09-949-016-12383	Sequence 12383, A
C 871	17	40.5	2184	3	US-09-540-236-1103	Sequence 1103, Ap	C 944	17	40.5	31063	3	US-09-596-002-20	Sequence 20, Appli
C 872	17	40.5	2197	3	US-08-985-950-13	Sequence 13, Appli	C 945	17	40.5	31166	3	US-09-949-016-13454	Sequence 13454, A
C 873	17	40.5	2197	3	US-09-546-049-13	Sequence 13, Appli	C 946	17	40.5	36090	3	US-09-949-016-17454	Sequence 17454, A
C 874	17	40.5	2200	3	US-08-985-950-19	Sequence 19, Appli	C 947	17	40.5	36123	3	US-09-949-016-12402	Sequence 12402, A
C 875	17	40.5	2200	3	US-09-546-049-19	Sequence 19, Appli	C 948	17	40.5	45225	3	US-09-949-016-12428	Sequence 12428, A
C 876	17	40.5	2211	3	US-09-107-532A-2353	Sequence 2353, Ap	C 949	17	40.5	45226	3	US-09-949-016-13654	Sequence 13654, A
C 877	17	40.5	2221	3	US-09-310-463-9	Sequence 9, Appli	C 950	17	40.5	45842	3	US-09-949-016-16914	Sequence 16914, A
C 878	17	40.5	2221	3	US-08-842-248A-9	Sequence 9, Appli	C 951	17	40.5	51508	3	US-09-949-016-16681	Sequence 16681, A
C 879	17	40.5	2221	5	US-10-143-618-9	Sequence 9, Appli	C 952	17	40.5	56523	3	US-09-949-016-14287	Sequence 14287, A
C 880	17	40.5	2271	3	US-08-985-950-15	Sequence 15, Appli	C 953	17	40.5	58361	3	US-09-949-016-16755	Sequence 16755, A
C 881	17	40.5	2271	3	US-09-546-049-15	Sequence 15, Appli	C 954	17	40.5	58361	3	US-09-949-016-16756	Sequence 16756, A
C 882	17	40.5	2271	3	US-09-252-991A-14238	Sequence 14238, A	C 955	17	40.5	59065	3	US-09-813-817-3	Sequence 3, Appli
C 883	17	40.5	2291	5	US-09-974-300-6330	Sequence 6330, Ap	C 956	17	40.5	59065	3	US-09-978-197-3	Sequence 3, Appli
C 884	17	40.5	2388	3	US-08-985-950-17	Sequence 17, Appli	C 957	17	40.5	59065	3	US-10-135-696-3	Sequence 3, Appli
C 885	17	40.5	2388	3	US-09-546-049-17	Sequence 17, Appli	C 958	17	40.5	59065	3	US-10-820-230-3	Sequence 3, Appli
C 886	17	40.5	2462	3	US-10-104-047-78	Sequence 78, Appli	C 959	17	40.5	59519	3	US-09-949-016-13504	Sequence 13504, A
C 887	17	40.5	2481	3	US-08-899-578-1	Sequence 1, Appli	C 960	17	40.5	60489	3	US-09-949-016-16287	Sequence 16287, A
C 888	17	40.5	2574	3	US-09-248-796A-6419	Sequence 6419, Ap	C 961	17	40.5	61042	3	US-09-949-016-13224	Sequence 13224, A
C 889	17	40.5	2576	4	US-10-094-749-1045	Sequence 1045, Ap	C 962	17	40.5	61735	3	US-09-949-016-12064	Sequence 12064, A
C 890	17	40.5	2724	3	US-10-104-047-1127	Sequence 1127, Ap	C 963	17	40.5	67620	3	US-09-949-016-16939	Sequence 16939, A
C 891	17	40.5	2860	2	US-08-462-484-7	Sequence 7, Appli	C 964	17	40.5	67911	3	US-09-949-016-16979	Sequence 16979, A
C 892	17	40.5	2860	2	US-08-441-147-7	Sequence 7, Appli	C 965	17	40.5	71815	3	US-09-949-016-12501	Sequence 12501, A
C 893	17	40.5	2860	7	PCT-US95-07536-7	Sequence 7, Appli	C 966	17	40.5	74527	3	US-09-949-016-12339	Sequence 12339, A
C 894	17	40.5	2909	3	US-10-104-047-614	Sequence 614, App	C 967	17	40.5	74528	3	US-09-949-016-13275	Sequence 13275, A
C 895	17	40.5	3001	3	US-09-539-333D-181	Sequence 181, App	C 968	17	40.5	82596	4	US-09-531-120-207	Sequence 207, App
C 896	17	40.5	3001	3	US-09-539-333D-197	Sequence 197, App	C 969	17	40.5	86439	3	US-09-949-016-11945	Sequence 11945, A
C 897	17	40.5	3078	3	US-09-623-551-28	Sequence 28, Appli	C 970	17	40.5	86440	3	US-09-949-016-16990	Sequence 16990, A
C 898	17	40.5	3092	3	US-09-566-921-121	Sequence 121, App	C 971	17	40.5	86980	3	US-09-949-016-15344	Sequence 15344, A
C 899	17	40.5	3134	2	US-07-865-662F-7	Sequence 7, Appli	C 972	17	40.5	92584	4	US-09-531-120-194	Sequence 194, App

973 17 40.5 94873 3 US-09-949-016-14277 Sequence 14277, A
974 17 40.5 94905 4 US-09-531-120-208 Sequence 208, App
975 17 40.5 98864 3 US-09-949-016-15403 Sequence 15403, A
976 17 40.5 99304 3 US-09-949-016-15440 Sequence 15440, A
977 17 40.5 103750 3 US-09-949-016-13319 Sequence 13319, A
978 17 40.5 105055 3 US-09-949-016-14001 Sequence 14001, A
979 17 40.5 105733 3 US-09-949-016-13080 Sequence 13080, A
980 17 40.5 118067 3 US-09-497-855A-32 Sequence 32, Appl
981 17 40.5 124110 3 US-09-949-016-13353 Sequence 13353, A
982 17 40.5 134499 4 US-09-531-120-192 Sequence 192, App
983 17 40.5 135010 3 US-09-949-016-17234 Sequence 17234, A
984 17 40.5 139250 3 US-09-949-016-17398 Sequence 17398, A
985 17 40.5 139577 3 US-09-949-016-12879 Sequence 12879, A
986 17 40.5 141589 5 US-09-543-679A-2480 Sequence 2480, Ap
987 17 40.5 141589 5 US-09-543-679A-2694 Sequence 2694, Ap
988 17 40.5 141589 5 US-09-543-679A-2719 Sequence 2719, Ap
989 17 40.5 145241 3 US-09-949-016-17394 Sequence 17394, A
990 17 40.5 145241 3 US-09-949-016-17395 Sequence 17395, A
991 17 40.5 145941 5 US-09-543-679A-3009 Sequence 3009, Ap
992 17 40.5 148794 3 US-09-949-016-12751 Sequence 12751, A
993 17 40.5 149543 3 US-09-949-016-15947 Sequence 15947, A
994 17 40.5 150597 3 US-09-949-016-15379 Sequence 15379, A
995 17 40.5 151828 4 US-09-531-120-197 Sequence 197, App
996 17 40.5 152524 3 US-09-949-016-12683 Sequence 12683, A
997 17 40.5 152524 3 US-09-949-016-13194 Sequence 13194, A
998 17 40.5 168174 3 US-10-071-411A-63 Sequence 63, Appl
999 17 40.5 168273 3 US-10-071-411A-2 Sequence 2, Appli
c1000 17 40.5 177293 3 US-09-949-016-16513 Sequence 16513, A

ALIGNMENTS

RESULT 1
US-09-949-016-29173
; Sequence 29173, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29173
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29173
Query Match 100.0%; Score 42; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 7.8e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 14 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 55
RESULT 2
US-09-949-016-126957
; Sequence 126957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126957
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126957

Query Match 100.0%; Score 42; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 7.8e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 14 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 55

RESULT 3

US-09-188-930-249
; Sequence 249, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-188-930-249

Query Match 100.0%; Score 42; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 9.4e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 990 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 4

US-09-312-283C-249
; Sequence 249, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-249

Query Match 100.0%; Score 42; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 9.4e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 990 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1031

RESULT 5

US-09-495-050A-223

; Sequence 223, Application US/09495050A

; Patent No. 6492505

; GENERAL INFORMATION:

; APPLICANT: Roopa, Reddy

; APPLICANT: Guegler, Karl, J.

; APPLICANT: Au-Young, Janice

; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P

; FILE REFERENCE: PA-0013 US

; CURRENT APPLICATION NUMBER: US/09/495,050A

; CURRENT FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/118,318

; PRIOR FILING DATE: February 1, 1999

; NUMBER OF SEQ ID NOS: 305

; SOFTWARE: PERL Program

; SEQ ID NO 223

; LENGTH: 1228

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6492505 2681738CT1

US-09-495-050A-223

Query Match 100.0%; Score 42; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 541 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 582

RESULT 6

US-09-949-016-3613

; Sequence 3613, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3613
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3613

Query Match 100.0%; Score 42; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1117 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 7

US-09-949-016-572

; Sequence 572, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 572

; LENGTH: 2456

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-572

Query Match 100.0%; Score 42; DB 3; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 8

US-09-949-016-15355

; Sequence 15355, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15355

; LENGTH: 26086

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-15355

```
Query Match      100.0%; Score 42; DB 3; Length 26086;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 22913 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 22954

RESULT 9
US-09-949-016-12314
; Sequence 12314, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12314
; LENGTH: 26238
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12314

Query Match      100.0%; Score 42; DB 3; Length 26238;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 23067 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 23108

RESULT 10
US-09-555-704-3
; Sequence 3, Application US/09555704
; Patent No. 6841374
; GENERAL INFORMATION:
; APPLICANT: The Research Foundation for Microbial diseases of Osaka University
; TITLE OF INVENTION: Japanese Encephalitis
; FILE REFERENCE: P98AF308-2
; CURRENT APPLICATION NUMBER: US/09/555,704
; CURRENT FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Japanese encephalitis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1500)
US-09-555-704-3

Query Match      54.3%; Score 22.8; DB 3; Length 1500;
Best Local Similarity 79.4%; Pred. No. 3.6;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCCGAGCCCTTACAAGACTATGAAG 37
Db 1203 CACGCTGGCGAAGCCCTTTCAACGACTTTGAAG 1236
```

```
RESULT 11
US-08-224-391-52
; Sequence 52, Application US/08224391
; Patent No. 5744140
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: PINCUS, STEVEN E.
; TITLE OF INVENTION: FLAVIVIRUS RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/224,391
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/729,800
; FILING DATE: 17-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4512 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-224-391-52

Query Match      54.3%; Score 22.8; DB 2; Length 4512;
Best Local Similarity 79.4%; Pred. No. 4.9;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CACGCTTGGCCGAGCCCTTACAAGACTATGAAG 37
Db 2085 CACGCTGGCGAAGCCCTTTCAACGACTTTGAAG 2118

RESULT 12
US-08-484-304-52
; Sequence 52, Application US/08484304
; Patent No. 5744141
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Pincus, Steven E.
; TITLE OF INVENTION: FLAVIVIRUS RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

RESULT 14 -
US-09-513-999C-20939/c
: Sequence 20939, Application US/09513999C

RESULT 16
US-09-949-016-177623
; Sequence 177623, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 177623
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-177623

Query Match 51.4%; Score 21.6; DB 3; Length 601;
Best Local Similarity 75.0%; Pred. No. 9.7;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAGACTATGAA 36
| | | | | | | | | | | | | | | | | | | | | |
Db 467 GTCCAGGCTTGTCAGCCCTTAGAAGGGCTAGCAA 502

RESULT 17
US-09-851-896-3/c
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851,896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match 51.4%; Score 21.6; DB 3; Length 70000;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAGACTATGAA 36
| | | | | | | | | | | | | | | | | | | | | |
Db 42352 GTCCAGGCTTGTCAGCCCTTAGAAGGGCTAGCAA 42317

RESULT 18
US-09-949-016-16819/c
; Sequence 16819, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16819
; LENGTH: 76399
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(76399)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16819

Query Match 51.4%; Score 21.6; DB 3; Length 76399;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAGACTATGAA 36
| | | | | | | | | | | | | | | | | | | | | |
Db 45406 GTCCAGGCTTGTCAGCCCTTAGAAGGGCTAGCAA 45371

RESULT 19
US-09-513-999C-32761
; Sequence 32761, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 32761
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 384
; OTHER INFORMATION: r=a or g
US-09-513-999C-32761

Query Match 51.0%; Score 21.4; DB 3; Length 407;
Best Local Similarity 80.6%; Pred. No. 11;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 11 GGCCGAGCCCTTACAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | | |
Db 10 GGCATGGCCTTCACATGACTATGAAGTAAA 40

RESULT 20
US-09-949-016-12756/c
; Sequence 12756, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.


```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1950
; LENGTH: 2772
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1950

Query Match
Best Local Similarity 49.0%; Score 20.6; DB 3; Length 2772;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 CTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1811 CTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1845

RESULT 26
US-09-949-016-13692
; Sequence 13692, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13692
; LENGTH: 36895
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13692

Query Match
Best Local Similarity 49.0%; Score 20.6; DB 3; Length 36895;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 CTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 33934 CTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 33968

RESULT 27
US-09-702-705-823/c
; Sequence 823, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13276
; LENGTH: 100877
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(100877)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13276

Query Match
Best Local Similarity 49.5%; Score 20.8; DB 3; Length 100877;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAA 40
Db 18788 GACCACGCTAGACAGTGCAGGAACAAAACATATATGTAA 18749

RESULT 24
US-09-949-016-17074/c
; Sequence 17074, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17074
; LENGTH: 140725
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17074

Query Match
Best Local Similarity 49.5%; Score 20.8; DB 3; Length 140725;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 29158 CCACGCCGAGCCCTTCTTCTAGAGCTGTAGTAAAG 29119

RESULT 25
US-09-949-016-1950
; Sequence 1950, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```



```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 823
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)..(264)
; OTHER INFORMATION: n=A,T,C or G
US-09-702-705-823

Query Match      48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 GCCCGGTNTGGCCNANCGCTGGCAAAGGCCCCAGNAGGAAAG 151

RESULT 28
US-09-736-457-823/c
; Sequence 823, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 823
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)..(264)
; OTHER INFORMATION: n=A,T,C or G
US-09-736-457-823

Query Match      48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 GCCCGGTNTGGCCNANCGCTGGCAAAGGCCCCAGNAGGAAAG 151

RESULT 29
US-09-614-124B-823/c
; Sequence 823, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
```

```
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 823
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)..(264)
; OTHER INFORMATION: n=A,T,C or G
US-09-614-124B-823

Query Match      48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 GCCCGGTNTGGCCNANCGCTGGCAAAGGCCCCAGNAGGAAAG 151

RESULT 30
US-09-671-325-823/c
; Sequence 823, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 823
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)..(264)
; OTHER INFORMATION: n=A,T,C or G
US-09-671-325-823

Query Match      48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 GCCCGGTNTGGCCNANCGCTGGCAAAGGCCCCAGNAGGAAAG 151

RESULT 31
US-09-589-184-823/c
; Sequence 823, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
```

APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C8
CURRENT APPLICATION NUMBER: US/09/589,184
CURRENT FILING DATE: 2000-06-05
NUMBER OF SEQ ID NOS: 827
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 823
LENGTH: 264
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(264)
OTHER INFORMATION: n=A,T,C or G
US-09-589-184-823

Query Match 48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 192 GCCCCGTTGGCCNANCGCTGGCAAAGGCCAGNAGGAAAG 151

RESULT 32
US-09-658-824-823/c
Sequence 823, Application US/09658824
Patent No. 6746846
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C11
CURRENT APPLICATION NUMBER: US/09/658,824
CURRENT FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 1788
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 823
LENGTH: 264
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(264)
OTHER INFORMATION: n=A,T,C or G
US-09-658-824-823

Query Match 48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 192 GCCCCGTTGGCCNANCGCTGGCAAAGGCCAGNAGGAAAG 151

RESULT 33
US-10-017-754-823/c
Sequence 823, Application US/10017754
Patent No. 6858204
GENERAL INFORMATION:
APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tongtong
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C18
CURRENT APPLICATION NUMBER: US/10/017,754
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 2004
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 823
LENGTH: 264
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 4, 7, 15, 17, 35, 38, 44, 53, 90, 105, 108, 115, 117,
LOCATION: 121, 126, 128, 158, 176, 178, 184, 201, 221, 227, 229, 233,
LOCATION: 239, 250
OTHER INFORMATION: n = A,T,C or G
US-10-017-754-823

Query Match 48.6%; Score 20.4; DB 3; Length 264;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 192 GCCCCGTTGGCCNANCGCTGGCAAAGGCCAGNAGGAAAG 151

RESULT 34
US-09-651-563-823/c
Sequence 823, Application US/09651563
Patent No. 6914132
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C10
CURRENT APPLICATION NUMBER: US/09/651,563
CURRENT FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 1679
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 823
LENGTH: 264
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(264)
OTHER INFORMATION: n=A,T,C or G

Qy 5 ACGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAG 42
 || || || || || || || || ||
Db 538 ACCGTTTGCAAGCAGATAAAAAGATGATCAAGTAAG 501

```

RESULT 38
US-10-094-749-1277
; Sequence 1277, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYOKU
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1277
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1277

```

```
; Query Match 48.6%; Score 20.4; DB 4; Length 2523;
; Best Local Similarity 71.1%; Pred. No. 49;
; Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGT 38
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 GCCCTTGGCGGCTGCGCCCTTACACAGACTATGTGGT 222

RESULT 39
US-10-104-047-349
; Sequence 349, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: H1-A0105
; FILE REFERENCE: No. 6943241el full length cDNA
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 349
; LENGTH: 2648
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-349

Query Match 48.6%; Score 20.4; DB 3; Length 2648;
; Best Local Similarity 71.1%; Pred. No. 50;
; Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGT 38
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 659 GCCCTTGGCGGCTGCGCCCTTACACAGACTATGTGGT 696

RESULT 40
US-09-902-540-1194
; Sequence 1194, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1194
; LENGTH: 21295
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1194

Query Match 48.6%; Score 20.4; DB 3; Length 21295;
; Best Local Similarity 71.1%; Pred. No. 89;
; Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGT 38
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15744 GCCCGAGCCTGGCTGACGCTTTCACAGCCTATGACGT 15781

RESULT 41
US-08-086-428B-49
; Sequence 49, Application US/08086428B
```

```
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
US-08-086-428B-49

Query Match 47.6%; Score 20; DB 2; Length 576;
; Best Local Similarity 72.2%; Pred. No. 50;
; Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCGAGCCCTTACAAGACTATGAAG 37
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 38 CCAATGATTGCCGAACTCTTCATAGTCTATGAGG 73

RESULT 42
US-08-468-570-49
; Sequence 49, Application US/08468570
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
```

ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,570
FILING DATE: 6-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA7
US-08-468-570-49

Query Match 47.6%; Score 20; DB 2; Length 576;
Best Local Similarity 72.2%; Pred. No. 50;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 CCACGCTTGGCGGAGCCCTTACAAAGACTATGAAG 37
DB 38 CCAATGATTCGCCGAACTCTTCATAGTCTATGAGG 73

RESULT 43
US-08-290-665A-49
Sequence 49, Application US/08290665A
Patent No. 582852
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK

REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA7
US-08-290-665A-49

Query Match 47.6%; Score 20; DB 2; Length 576;
Best Local Similarity 72.2%; Pred. No. 50;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 CCACGCTTGGCGGAGCCCTTACAAAGACTATGAAG 37
DB 38 CCAATGATTCGCCGAACTCTTCATAGTCTATGAGG 73

RESULT 44
US-08-466-601A-49
Sequence 49, Application US/08466601A
Patent No. 6572864
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
OF REAGENTS DERIVED FROM THESE SEQUENCES IN
DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,601A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
US-08-466-601A-49

Query Match          47.6%; Score 20; DB 3; Length 576;
Best Local Similarity 72.2%; Pred. No. 50;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CCACGCTTGGCGGAGCCCTTACAAAGACTATGAAG 37
Db 38 CCAATGATTGCCGGAACCTTCCATAGTCTATGAGG 73

RESULT 45
PCT-US95-10398-49
; Sequence 49, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
PCT-US95-10398-49

Query Match          47.6%; Score 20; DB 7; Length 576;
Best Local Similarity 72.2%; Pred. No. 50;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CCACGCTTGGCGGAGCCCTTACAAAGACTATGAAG 37
Db 38 CCAATGATTGCCGGAACCTTCCATAGTCTATGAGG 73

us-10-600-816-30.rni

; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
US-08-466-601A-49

Query Match          47.6%; Score 20; DB 3; Length 576;
Best Local Similarity 72.2%; Pred. No. 50;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CCACGCTTGGCGGAGCCCTTACAAAGACTATGAAG 37
Db 38 CCAATGATTGCCGGAACCTTCCATAGTCTATGAGG 73

RESULT 46
US-09-949-016-15770
; Sequence 15770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15770
; LENGTH: 211049
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-15770

Query Match          47.6%; Score 20; DB 3; Length 211049;
Best Local Similarity 82.1%; Pred. No. 2.5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGGAGCCCTTACAAAG 28
Db 113886 GCCACGCTTGGCGGAGCCCTTACAAAG 113913

RESULT 47
US-09-533-559-647/c
; Sequence 647, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 647
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(546)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-647

Query Match          47.1%; Score 19.8; DB 3; Length 546;
Best Local Similarity 67.5%; Pred. No. 60;
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAA 40
Db 113886 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAA 40
```



```
Db 491 GCCCTTGCTAGAGGCGTTGACAAAGTCCAGGAAGTCA 452

RESULT 48
US-09-949-016-148482/c
; Sequence 148482, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148482
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-148482

Query Match 47.1%; Score 19.8; DB 3; Length 601;
Best Local Similarity 77.4%; Pred. No. 62;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCGCCGAGCCCTTACAAAGACTA 32
Db 524 CCCACCTGGCCCAAGCCCAACTCAGACTA 494

RESULT 49
US-09-949-016-148483/c
; Sequence 148483, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148483
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-148483

Query Match 47.1%; Score 19.8; DB 3; Length 601;
Best Local Similarity 77.4%; Pred. No. 62;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCGCCGAGCCCTTACAAAGACTA 32
Db 260 CCCACCTGGCCCAAGCCCAACTCAGACTA 230

RESULT 50
US-09-717-432-1/c
```

```
; Sequence 1, Application US/09717432
; Patent No. 6291223
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: LI, XIAOTONG
; APPLICANT: CHRISTIE, GARY
; APPLICANT: POWELL, DAVID J.
; TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASp1)
; FILE REFERENCE: GP-70663
; CURRENT APPLICATION NUMBER: US/09/717,432
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/166,974
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: MUS MUSCULUS
US-09-717-432-1

Query Match 47.1%; Score 19.8; DB 3; Length 1545;
Best Local Similarity 69.2%; Pred. No. 80;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCCGAGCCCTTACAAAGACTATGAAGTA 39
Db 1269 GCCACCCCTCTGTGAGCTCTGTCAAAGACCACGTAGAA 1231

RESULT 51
US-09-912-484-1/c
; Sequence 1, Application US/09912484
; Patent No. 6358725
; GENERAL INFORMATION:
; APPLICANT: Christie, Gary
; APPLICANT: Li, Xiaotong
; APPLICANT: Powell, David J.
; APPLICANT: Zhu, Yuan
; TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASp1)
; FILE REFERENCE: GP-70663-D1
; CURRENT APPLICATION NUMBER: US/09/912,484
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/166,974
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/717,432
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: MUS MUSCULUS
US-09-912-484-1

Query Match 47.1%; Score 19.8; DB 3; Length 1545;
Best Local Similarity 69.2%; Pred. No. 80;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCCGAGCCCTTACAAAGACTATGAAGTA 39
Db 1269 GCCACCCCTCTGTGAGCTCTGTCAAAGACCACGTAGAA 1231

RESULT 52
US-09-949-016-15910/c
; Sequence 15910, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15910
; LENGTH: 8961
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15910

Query Match 47.1%; Score 19.8; DB 3; Length 8961;
Best Local Similarity 77.4%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCCACGCTTGGCGAGCCCTTACAAAGACTA 32
||||| ||||| ||||| ||||| |||||
Db 5129 CCCACCTGGCCGAGCCCAACTCAGACTA 5099

RESULT 53

US-09-949-016-15306/c
; Sequence 15306, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15306
; LENGTH: 9382
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15306

Query Match 47.1%; Score 19.8; DB 3; Length 9382;
Best Local Similarity 77.4%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCCACGCTTGGCGAGCCCTTACAAAGACTA 32
||||| ||||| ||||| ||||| |||||
Db 1556 CCCAGGCTGCCCTAGCCTTGCATAATA 1526

RESULT 54

US-09-949-016-13313/c
; Sequence 13313, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13313
; LENGTH: 75212
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13313

Query Match 47.1%; Score 19.8; DB 3; Length 75212;
Best Local Similarity 69.2%; Pred. No. 2.3e+02;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
||||| ||||| ||||| ||||| |||||
Db 29029 CCATGCTGGCCGACACTATAATTTCTATTATTGAA 28991

RESULT 55

US-09-949-016-13314/c
; Sequence 13314, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13314
; LENGTH: 75212
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13314

Query Match 47.1%; Score 19.8; DB 3; Length 75212;
Best Local Similarity 69.2%; Pred. No. 2.3e+02;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
||||| ||||| ||||| ||||| |||||
Db 29029 CCATGCTGGCCGACACTATAATTTCTATTATTGAA 28991

RESULT 56

US-09-949-016-13315/c
; Sequence 13315, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0


```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14083
; LENGTH: 199471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(199471)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14083

Query Match      47.1%; Score 19.8; DB 3; Length 199471;
Best Local Similarity 69.2%; Pred. No. 3e+02;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2 CCCACGTTGCCGAGCCCTTACAAAGACTATGAAGTAA 40
Db 100164 CCCACCTACTCTCTGACCTGACTGAGAGTGTGAATGA 100202

RESULT 66
US-08-956-171E-4323
; Sequence 4323, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
```

```
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 4323:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4323:
US-08-956-171E-4323

Query Match      46.7%; Score 19.6; DB 3; Length 337;
Best Local Similarity 73.5%; Pred. No. 64;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 TTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 160 TTGCGCGGCCCCCAACAAGAGCTGACGAAAG 193

RESULT 67
US-08-781-986A-4323
; Sequence 4323, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 4323:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-4323

Query Match      46.7%; Score 19.6; DB 3; Length 337;
Best Local Similarity 73.5%; Pred. No. 64;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 TTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 160 TTGCGCGGCCCCCAACAAGAGCTGACGAAAG 193

RESULT 68
US-09-949-016-146770/c
; Sequence 146770, Application US/09949016
```

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146770
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-146770

Query Match          46.7%; Score 19.6; DB 3; Length 601;
Best Local Similarity 73.5%; Pred. No. 76;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
   ||||| ||| ||| ||| ||| ||| ||| |||
DB 369 TTGCCGAGCTTTTCATAAACTTGAATATAG 336

RESULT 69
US-09-205-815B-49/c
; Sequence 49, Application US/09205815B
; Patent No. 6492509
; GENERAL INFORMATION:
; APPLICANT: Calgene, LLC
; TITLE OF INVENTION: Fatty Acyl-CoA: Fatty Alcohol Acyltransferases
; FILE REFERENCE: Fatty Acyl-CoA: Fatty Alcohol Acyltran
; CURRENT APPLICATION NUMBER: US/09/205,815B
; CURRENT FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: 09/092,562
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,651
; PRIOR FILING DATE: 1997-06-05
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: IBM PC; Windows NT 4.0; Microsoft Word For Windows 7.0a
; SEQ ID NO 49
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: Gene
; LOCATION: (1)..(1020)
; OTHER INFORMATION: sequence of AT_WS7
; US-09-205-815B-49

Query Match          46.7%; Score 19.6; DB 3; Length 1020;
Best Local Similarity 73.5%; Pred. No. 87;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 5 ACCTTGGCGAGCCCTTACAAAGACTATGAAGT 38
   ||||| ||| ||| ||| ||| ||| |||
DB 623 ACGATTGCCGAGAGACTTGGCGAGACTATGAGT 590

RESULT 70
5486473-1
; Patent No. 5486473
; APPLICANT: FUKITA, HIROYUKI; YOSHIDA, IWA0; TAKAGI, MITSUO;
; MANABE, SADA0; FUKAI, KONOSUKE
; TITLE OF INVENTION: A DNA CODING FOR A FLAVIVIRUS ANTIGEN
; . . . NUMBER OF SEQUENCES: 16
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,049
; FILING DATE: 09-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 809,255
; FILING DATE: 18-DEC-1991
; APPLICATION NUMBER: 279,685
; FILING DATE: 05-DEC-1988
; APPLICATION NUMBER: 932,419
; FILING DATE: 19-NOV-1986
; SEQ ID NO:1:
; LENGTH: 1500
5486473-1

Query Match          46.7%; Score 19.6; DB 10; Length 1500;
Best Local Similarity 73.5%; Pred. No. 97;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
   ||||| ||| ||| ||| ||| ||| |||
DB 1203 CACGCTAGGCAAGGCGCTTTTCAACAACTTTGAAG 1236

RESULT 71
US-08-956-171E-316
; Sequence 316, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELCOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 316:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 316:
US-08-956-171E-316
```


Query Match 46.7%; Score 19.6; DB 3; Length 9310;
Best Local Similarity 73.5%; Pred. No. 1.6e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 TTGCGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3885 TTGCGGGGGCCCCAACAAAGAGCTGACGAAAG 3918

RESULT 72

US-08-781-986A-316
; Sequence 316, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 316:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-316

Query Match 46.7%; Score 19.6; DB 3; Length 9310;
Best Local Similarity 73.5%; Pred. No. 1.6e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 TTGCGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3885 TTGCGGGGGCCCCAACAAAGAGCTGACGAAAG 3918

RESULT 73

US-09-452-638-52
; Sequence 52, Application US/09452638
; Patent No. 6696281
; GENERAL INFORMATION:
; APPLICANT: Chambers, Thomas J.
; APPLICANT: Monath, Thomas P.
; APPLICANT: Guirakhoo, Farshad
; TITLE OF INVENTION: Chimeric Flavivirus Vaccines
; FILE REFERENCE: 06132/033004
; CURRENT APPLICATION NUMBER: US/09/452,638
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 09/121,587

; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/US98/03894
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: US 09/007,664
; PRIOR FILING DATE: 1998-01-15
; PRIOR APPLICATION NUMBER: US 08/807,445
; PRIOR FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 10892
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: derived from Yellow Fever virus and Japanese
; OTHER INFORMATION: Encephalitis virus
; NAME/KEY: CDS
; LOCATION: (119)...(10381)
US-09-452-638-52

Query Match 46.7%; Score 19.6; DB 3; Length 10892;
Best Local Similarity 73.5%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 4 CACGCTTGCGCGAGCCCTTACAAAGACTATGAAG 37
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2185 CACGCTGGGCAAGGCCCTTTTCAACAACTTTGAAG 2218

RESULT 74

US-09-121-587A-12
; Sequence 12, Application US/09121587A
; Patent No. 6962708
; GENERAL INFORMATION:
; APPLICANT: Chambers, Thomas J.
; APPLICANT: Guirakhoo, Farshad
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: CHIMERIC FLAVIVIRUS VACCINES
; FILE REFERENCE: 06132/033003
; CURRENT APPLICATION NUMBER: US/09/121,587A
; CURRENT FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: US98/03894
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 09/007,664
; PRIOR FILING DATE: 1998-01-15
; PRIOR APPLICATION NUMBER: 08/807,445
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 10892
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: derived from Japanese Encephalitis virus and
; OTHER INFORMATION: Yellow Fever virus
; NAME/KEY: CDS
; LOCATION: (119)...(10381)
US-09-121-587A-12

Query Match 46.7%; Score 19.6; DB 3; Length 10892;
Best Local Similarity 73.5%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 4 CACGCTTGCGCGAGCCCTTACAAAGACTATGAAG 37
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2185 CACGCTGGGCAAGGCCCTTTTCAACAACTTTGAAG 2218

RESULT 75

US-09-949-002-832
; Sequence 832, Application US/09949002
; Patent No. 6900016

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 832
; LENGTH: 36677
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-832

Query Match          46.7%; Score 19.6; DB 3; Length 36677;
Best Local Similarity 73.5%; Pred. No. 2.4e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18546 TGGCGAGAGCCCTTAAGGAACAATGAAGTAAAG 18579.

RESULT 76
US-09-949-016-15863
; Sequence 15863, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15863
; LENGTH: 96340
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15863

Query Match          46.7%; Score 19.6; DB 3; Length 96340;
Best Local Similarity 73.5%; Pred. No. 3.1e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGCGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36672 TTGCGCCAGTCCTTTTCATAAATTTGAATATAG 36705

RESULT 77
US-09-949-016-13750/c
; Sequence 13750, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13750
; LENGTH: 132456
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(132456)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13750

Query Match          46.7%; Score 19.6; DB 3; Length 132456;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85965 GCCCCGCTCCGCCCGTCCATAAAATGGCTAGGCTGTCAAG 85924

RESULT 78
US-09-949-016-15436/c
; Sequence 15436, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15436
; LENGTH: 199945
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15436

Query Match          46.7%; Score 19.6; DB 3; Length 199945;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134050 GCCACTGTGGGCCCGCCGCTTAGAGAGGCCCTGAAGCCCGAG 134009

RESULT 79
US-09-949-016-14588
; Sequence 14588, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14588
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14588

Query Match          46.7%; Score 19.6; DB 3; Length 302604;
Best Local Similarity 73.5%; Pred. No. 4.1e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 7 GCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 40
Db 1278 GCTTCTGTGAGTCTCTACAAGACTCCGGAGTAA 1311

RESULT 80
US-09-949-016-14589
; Sequence 14589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14589
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14589

Query Match          46.7%; Score 19.6; DB 3; Length 302604;
Best Local Similarity 73.5%; Pred. No. 4.1e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 7 GCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 40
Db 1278 GCTTCTGTGAGTCTCTACAAGACTCCGGAGTAA 1311

RESULT 81
US-09-949-016-17119
; Sequence 17119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17119
; LENGTH: 308362
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(308362)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17119

Query Match          46.7%; Score 19.6; DB 3; Length 308362;
Best Local Similarity 73.5%; Pred. No. 4.1e+02;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 7 GCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 40
Db 1094 GCTTCTGTGAGTCTCTACAAGACTCCGGAGTAA 1127

RESULT 82
US-09-621-976-16747
; Sequence 16747, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16747
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16747

Query Match          46.2%; Score 19.4; DB 3; Length 461;
Best Local Similarity 63.4%; Pred. No. 86;
Matches 26; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAGCCCTTACAAGACTATGAAGTAA 41
Db 367 GCCCATGAGAGCGTGTCCTTTACAAGGAGTCTCARTGATA 407

RESULT 83
US-09-949-016-95958/c
; Sequence 95958, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; CURRENT APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95958
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-95958

Query Match          46.2%; Score 19.4; DB 3; Length 601;
Best Local Similarity 70.3%; Pred. No. 93;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 5 ACCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 41
Db 378 ACTCTGGCCAGTAAAGACAGCGACTATGAAAAA 342

RESULT 84
US-09-949-016-188486
; Sequence 188486, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188486
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-188486

Query Match          46.2%; Score 19.4; DB 3; Length 601;
Best Local Similarity 70.3%; Pred. No. 93;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 6 CGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 367 CACTTGGCCAATCCATGACAAAGTCCATGAACA 403

RESULT 85
US-09-270-767-12087/c
; Sequence 12087, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12087
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12087

Query Match          46.2%; Score 19.4; DB 3; Length 938;
Best Local Similarity 70.3%; Pred. No. 1e+02;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 6 CGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 141707 CACTTGGCCAATCCATGACAAAGTCCATGA 141743

Query Match          46.2%; Score 19.4; DB 3; Length 158735;
Best Local Similarity 70.3%; Pred. No. 4.3e+02;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 6 CGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 141707 CACTTGGCCAATCCATGACAAAGTCCATGA 141743

Query Match          46.2%; Score 19.4; DB 3; Length 158735;
Best Local Similarity 70.3%; Pred. No. 4.3e+02;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 6 CGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 141707 CACTTGGCCAATCCATGACAAAGTCCATGA 141743
```

RESULT 88
US-09-949-016-14543
; Sequence 14543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14543
; LENGTH: 387902
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(387902)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14543
Query Match 46.2%; Score 19.4; DB 3; Length 387902;
Best Local Similarity 70.3%; Pred. No. 5.3e+02;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 5 ACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 41
Db 28249 ACTCTGGCCAGTAAAGACGCGACTATGAAAAA 28285
RESULT 89
US-09-949-016-12557
; Sequence 12557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12557
; LENGTH: 421883
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(421883)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12557
Query Match 46.2%; Score 19.4; DB 3; Length 421883;
Best Local Similarity 70.3%; Pred. No. 5.3e+02;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 5 ACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 41

Db 28249 ACTCTGGCCAGTAAAGACGCGACTATGAAAAA 28285
RESULT 90
US-09-621-976-16083
; Sequence 16083, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16083
; LENGTH: 215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16083
Query Match 45.7%; Score 19.2; DB 3; Length 215;
Best Local Similarity 75.0%; Pred. No. 86;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 9 TTGCGCGAGCCCTTACAAGACTATGAAGTAA 40
Db 118 TTGCGCAAAACCTTCAAGGATATGCAGAA 149
RESULT 91
US-09-252-991A-8407
; Sequence 8407, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8407
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8407
Query Match 45.7%; Score 19.2; DB 3; Length 450;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 3 CCACGGTTGGCCGAGCCCTTACAAGACTATG 34
Db 107 CAAGCGCTGGCCGAGCCGCTCCATCGACTATG 138
RESULT 92
US-09-878-281A-153
; Sequence 153, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35


```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47026
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-47026

```

Query Match 45.7%; Score 19.2; DB 3; Length 601;
Best Local Similarity 67.5%; Pred. No. 1.1e+02;
Matches 27; Conservative 0; Mismatches 13; Indels

Qy 3 CCACGCTTGGCCGAGCCCTTACAAGA GACTATGAAGTAAG 42
| | | | | | | | | | | | | | | |
Db 50 CAAC TCTTGGGTTCCTGTCA CAGAGACTATAATGTGAAG 89

RESULT 98
US-08-612-973-31
; Sequence 31, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
;

ADDRESS: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON

```

/ STATE: VIRGINIA
/ COUNTRY: U.S.A.
/ ZIP: 22201-4714
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,973
/

```

FILING DATE: 11-MAR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 31:

```

1 INFORMATION FOR SEQ ID NO: 31
2
3 SEQUENCE CHARACTERISTICS:
4     LENGTH: 630 base pairs
5     TYPE: nucleic acid
6     STRANDEDNESS: single
7     TOPOLOGY: linear
8     MOLECULE TYPE: cDNA
9     HYPOTHETICAL: NO
10    ANTI-SENSE: NO
11
12    FEATURE:
13     NAME/KEY: CDS
14     LOCATION: 1..627

```

```

; LOCATION: 1..624
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624
US-08-612-973-31

```

Query Match	45.7%;	Score 19.2;	DB 3;	Length 630;
Best Local Similarity	67.5%;	Pred. No. 1.2e+02;		

0
aps

Qy 3 CCACGCTTGGCCGAGCCCTTACAAGA GACTATGAAGTAAG 42
| | | | | | | | | | | | | | | |
Db 50 CAAC TCTTGGGTTCCTGTCA CAGAGACTATAATGTGAAG 89

RESULT 98
US-08-612-973-31
; Sequence 31, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
;

ADDRESS: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON

```

/ STATE: VIRGINIA
/ COUNTRY: U.S.A.
/ ZIP: 22201-4714
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,973
/

```

FILING DATE: 11-MAR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 31:

```

1 INFORMATION FOR SEQ ID NO: 31
2
3 SEQUENCE CHARACTERISTICS:
4     LENGTH: 630 base pairs
5     TYPE: nucleic acid
6     STRANDEDNESS: single
7     TOPOLOGY: linear
8     MOLECULE TYPE: cDNA
9     HYPOTHETICAL: NO
10    ANTI-SENSE: NO
11
12    FEATURE:
13     NAME/KEY: CDS
14     LOCATION: 1..627

```

```

; LOCATION: 1..624
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624
US-08-612-973-31

```

Query Match	45.7%;	Score 19.2;	DB 3;	Length 630;
Best Local Similarity	67.5%;	Pred. No. 1.2e+02;		

```

Matches 27; Conservative 0; Mismatches 13; Indels 0;
Qy 2 CCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAA 41
Db 257 CCAATGATTGCCCAAACTCTTCCATAGTCTATGAGGAGA 296

RESULT 99
US-08-927-597-31
; Sequence 31, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (PFO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597

```

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:

PRIOR APPLICATION DATA: US 08/612,973
APPLICATION NUMBER: 11-MAR-1996
FILING DATE: 32,205
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 1487-10
REFERENCE/DOCKET NUMBER: 816-4000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100

```

? INFORMATION FOR SEQ ID NO: 31:
? =====
? LENGTH: 630 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FEATURE:

```

```

? NAME/KEY: CDS
? LOCATION: 1..627
? FEATURE:
?
? NAME/KEY: mat_peptide
? LOCATION: 1..624
?
US-08-927-597-31

Query Watch 45
Best Local Similarity 67
Matches 27; Conservative

```

Qy	2	CCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAA	41
Db	257	CCAATGATTGCCCAACCTTCCTAGTCTTGAGGCAGA	296

RESULT 100
US-08-928-757-31

; Sequence 31, Application US/08928757
; Patent No. 6890737
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; BOSMAN, FONS
; DE MARTYNOFF, GUY
; BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/08/928,757
; FILING DATE: 12-Sep-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..627
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-08-928-757-31

Query Match 45.7%; Score 19.2; DB 3; Length 630;
Best Local Similarity 67.5%; Pred. No. 1.2e+02;
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CCCACGCTGGCGGAGCCCTTACAAAGACTATGAAGTAA 41
Db 257 CCAATGATTGCCCAAACTCTTCATATGAGGAGA 296

Search completed: June 10, 2006, 15:37:12
Job time : 150.15 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:26:46 ; Search time 1608.6 Seconds
(without alignments)
320.826 Million cell updates/sec

Title: US-10-600-816-30

Perfect score: 42

Sequence: 1 gccacgctggccgagccc.....acaaagactgaagtaag 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications NA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10K_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	42	8	US-10-600-816-30
2	42	100.0	497	6	US-10-066-543-1937
3	42	100.0	552	6	US-10-066-543-1811
4	42	100.0	620	3	US-09-969-034-2222
5	42	100.0	634	3	US-09-969-034-3393
6	42	100.0	642	7	US-10-125-968-701
7	42	100.0	1071	3	US-10-712-615-134
8	42	100.0	1212	3	US-09-866-050A-249
9	42	100.0	1212	6	US-10-152-661-249
10	42	100.0	1228	7	US-10-313-542-223
11	42	100.0	1400	13	US-11-060-756-2418
12	42	100.0	1400	13	US-11-060-756-6690
13	42	100.0	1460	10	US-10-936-626-64
14	42	100.0	1460	10	US-10-938-061-64
15	42	100.0	1619	7	US-10-224-289-5
16	42	100.0	1619	9	US-10-935-190-43
17	42	100.0	2297	9	US-10-775-920-13

Sequence 3, Appli
Sequence 405, App
Sequence 2, Appli
Sequence 12, Appl
Sequence 59, Appl
Sequence 11, Appl
Sequence 45, App
Sequence 63, Appl
Sequence 64, Appl
Sequence 619, App
Sequence 2, Appli
Sequence 18, Appl
Sequence 10, Appl
Sequence 40, Appl
Sequence 32, Appl
Sequence 834, App
Sequence 10424, A
Sequence 196, App
Sequence 33, Appl
Sequence 43361, A
Sequence 7, Appli
Sequence 12, Appl
Sequence 9, Appli
Sequence 21, Appl
Sequence 691738,
Sequence 691739,
Sequence 691740,
Sequence 691738,
Sequence 691739,
Sequence 691740,
Sequence 76, Appl
Sequence 49510, A
Sequence 49510, A
Sequence 150748,
Sequence 764157,
Sequence 455, App
Sequence 409, App
Sequence 126, App
Sequence 304, App
Sequence 1315, Ap
Sequence 10139, A
Sequence 1115, Ap
Sequence 1131, Ap
Sequence 10138, A
Sequence 1313, Ap
Sequence 22855, A
Sequence 3, Appli
Sequence 1964, Ap
Sequence 26327, A
Sequence 28218, A
Sequence 36568, A
Sequence 130787,
Sequence 3621, Ap
Sequence 9335, Ap
Sequence 5769, Ap
Sequence 12052, A
Sequence 1493, Ap
Sequence 10499, A
Sequence 18436, A
Sequence 18787, A
Sequence 20633, A
Sequence 20810, A
Sequence 20790, A
Sequence 420, App
Sequence 5576, Ap
Sequence 214, App
Sequence 148760,
Sequence 148761,
Sequence 148760,
Sequence 148761,

C 91	20.8	49.5	569	12	US-10-301-480-242891	Sequence 242891, A	164	20.2	48.1	542	5	US-09-925-065A-579266	Sequence 579266,
C 92	20.8	49.5	569	12	US-10-301-480-242892	Sequence 242892, A	C 165	20.2	48.1	582	4	US-09-925-065A-491469	Sequence 491469,
C 93	20.8	49.5	569	12	US-10-301-480-856300	Sequence 856300, A	166	20.2	48.1	582	4	US-09-925-065A-505297	Sequence 505297,
C 94	20.8	49.5	569	12	US-10-301-480-856301	Sequence 856301, A	167	20.2	48.1	582	4	US-09-925-065A-505298	Sequence 505298,
C 95	20.8	49.5	1465	4	US-09-925-065A-716598	Sequence 716598, A	C 168	20.2	48.1	582	5	US-09-925-065A-491469	Sequence 491469,
C 96	20.8	49.5	1465	4	US-09-925-065A-716599	Sequence 716599, A	C 169	20.2	48.1	582	5	US-09-925-065A-505297	Sequence 505297,
C 97	20.8	49.5	1465	4	US-09-925-065A-716600	Sequence 716600, A	170	20.2	48.1	582	5	US-09-925-065A-505298	Sequence 505298,
C 98	20.8	49.5	1465	5	US-09-925-065A-716598	Sequence 716598, A	171	20.2	48.1	584	4	US-09-925-065A-505299	Sequence 505299,
C 99	20.8	49.5	1465	5	US-09-925-065A-716599	Sequence 716599, A	172	20.2	48.1	584	4	US-09-925-065A-505300	Sequence 505300,
C 100	20.8	49.5	1465	5	US-09-925-065A-716600	Sequence 716600, A	173	20.2	48.1	584	5	US-09-925-065A-505299	Sequence 505299,
C 101	20.6	49.0	170	8	US-10-242-535A-517625	Sequence 517625, A	174	20.2	48.1	584	5	US-09-925-065A-505300	Sequence 505300,
C 102	20.6	49.0	170	8	US-10-085-783A-51725	Sequence 51725, A	C 175	20.2	48.1	590	12	US-10-301-480-204736	Sequence 204736,
C 103	20.6	49.0	587	3	US-09-918-995-9395	Sequence 9395, Ap	C 176	20.2	48.1	590	12	US-10-301-480-204737	Sequence 204737,
C 104	20.6	49.0	584	4	US-09-925-065A-582213	Sequence 582213, A	C 177	20.2	48.1	590	12	US-10-301-480-818145	Sequence 818145,
C 105	20.6	49.0	584	4	US-09-925-065A-582214	Sequence 582214, A	C 178	20.2	48.1	590	12	US-10-301-480-818146	Sequence 818146,
C 106	20.6	49.0	584	5	US-09-925-065A-582213	Sequence 582213, A	C 179	20.2	48.1	594	4	US-09-925-065A-104585	Sequence 104585,
C 107	20.6	49.0	584	5	US-09-925-065A-582214	Sequence 582214, A	C 180	20.2	48.1	594	4	US-09-925-065A-104586	Sequence 104586,
C 108	20.6	49.0	593	9	US-10-425-115-128629	Sequence 128629, A	C 181	20.2	48.1	594	5	US-09-925-065A-104585	Sequence 104585,
C 109	20.6	49.0	600	10	US-10-956-157-10063	Sequence 10063, A	C 182	20.2	48.1	594	5	US-09-925-065A-104586	Sequence 104586,
C 110	20.6	49.0	649	10	US-10-956-157-3509	Sequence 3509, Ap	183	20.2	48.1	595	4	US-09-925-065A-543279	Sequence 543279,
C 111	20.6	49.0	649	10	US-10-956-157-3510	Sequence 3510, Ap	184	20.2	48.1	595	4	US-09-925-065A-543280	Sequence 543280,
C 112	20.6	49.0	649	10	US-10-956-157-8744	Sequence 8744, Ap	185	20.2	48.1	595	4	US-09-925-065A-543281	Sequence 543281,
C 113	20.6	49.0	649	10	US-10-956-157-8745	Sequence 8745, Ap	186	20.2	48.1	595	5	US-09-925-065A-543279	Sequence 543279,
C 114	20.6	49.0	682	3	US-09-814-353-19823	Sequence 19823, Ap	187	20.2	48.1	595	5	US-09-925-065A-543280	Sequence 543280,
C 115	20.6	49.0	934	8	US-10-425-114-31484	Sequence 31484, A	188	20.2	48.1	595	5	US-09-925-065A-543281	Sequence 543281,
C 116	20.6	49.0	966	12	US-10-301-480-601536	Sequence 601536, A	C 189	20.2	48.1	600	10	US-10-972-079-89493	Sequence 89493, A
C 117	20.6	49.0	966	12	US-10-301-480-1214945	Sequence 1214945, A	C 190	20.2	48.1	600	10	US-10-972-079-89494	Sequence 89494, A
C 118	20.6	49.0	1764	3	US-09-925-301-327	Sequence 327, App	C 191	20.2	48.1	609	4	US-09-925-065A-801931	Sequence 801931,
C 119	20.6	49.0	1818	16	US-11-096-568A-25665	Sequence 25665, A	C 192	20.2	48.1	609	5	US-09-925-065A-801931	Sequence 801931,
C 120	20.6	49.0	2085	9	US-10-357-930-22086	Sequence 22086, A	C 193	20.2	48.1	644	5	US-09-925-065A-365219	Sequence 365219,
C 121	20.6	49.0	2085	9	US-10-357-930-22183	Sequence 22183, A	194	20.2	48.1	644	5	US-09-925-065A-365219	Sequence 365219,
C 122	20.6	49.0	2085	9	US-10-357-930-27948	Sequence 27948, A	195	20.2	48.1	650	12	US-10-301-480-435794	Sequence 435794,
C 123	20.6	49.0	2085	9	US-10-357-930-28042	Sequence 28042, A	196	20.2	48.1	650	12	US-10-301-480-1049203	Sequence 1049203,
C 124	20.6	49.0	2286	9	US-10-425-115-128635	Sequence 128635, A	197	20.2	48.1	1410	8	US-10-282-122A-23586	Sequence 23586, A
C 125	20.6	49.0	4210	10	US-10-956-157-4828	Sequence 4828, Ap	198	20.2	48.1	1545	8	US-10-275-026A-85	Sequence 85, Appl
C 126	20.6	49.0	4522	10	US-10-450-763-668	Sequence 668, App	199	20.2	48.1	1581	10	US-10-750-185-64370	Sequence 64370, A
C 127	20.6	49.0	6579	6	US-10-128-714-21235	Sequence 21235, Ap	200	20.2	48.1	1581	10	US-10-750-623-64370	Sequence 64370, A
C 128	20.4	48.6	264	3	US-09-736-457-823	Sequence 823, App	201	20.2	48.1	6446	16	US-11-136-527-543	Sequence 543, App
C 129	20.4	48.6	264	3	US-09-902-941-823	Sequence 823, App	202	20.2	48.1	11851	10	US-10-661-398-24	Sequence 24, Appl
C 130	20.4	48.6	264	3	US-09-849-626-823	Sequence 823, App	C 203	20.2	48.1	108359	6	US-10-191-807-3	Sequence 3, Appl
C 131	20.4	48.6	264	6	US-10-017-754-823	Sequence 823, App	C 204	20.2	48.1	108359	10	US-10-959-243-3	Sequence 3, Appl
C 132	20.4	48.6	264	7	US-10-113-872-823	Sequence 823, App	C 205	20.2	48.1	187	8	US-10-437-963-95922	Sequence 95922, A
C 133	20.4	48.6	264	7	US-10-283-017-823	Sequence 823, App	C 206	20.2	47.6	389	9	US-10-425-115-6531	Sequence 6531, Ap
C 134	20.4	48.6	439	3	US-09-796-692-4516	Sequence 4516, Ap	207	20.2	47.6	489	8	US-10-021-323-11958	Sequence 11958, A
C 135	20.4	48.6	439	6	US-10-040-862-4516	Sequence 4516, Ap	208	20.2	47.6	531	8	US-10-258-368-9	Sequence 9, Appl
C 136	20.4	48.6	439	7	US-10-057-475B-4516	Sequence 4516, Ap	209	20.2	47.6	542	6	US-10-115-192-4	Sequence 4, Appl
C 137	20.4	48.6	439	7	US-10-154-884B-4516	Sequence 4516, Ap	210	20.2	47.6	546	8	US-10-258-368-16	Sequence 16, Appl
C 138	20.4	48.6	439	9	US-10-764-324-4516	Sequence 4516, Ap	C 211	20.2	47.6	608	3	US-10-021-323-16211	Sequence 16211, A
C 139	20.4	48.6	629	9	US-10-357-930-55715	Sequence 55715, A	212	20.2	47.6	649	3	US-09-814-353-16883	Sequence 16883, A
C 140	20.4	48.6	707	3	US-09-974-300-5294	Sequence 5294, Ap	213	20.2	47.6	667	9	US-10-767-795-3521	Sequence 3521, Ap
C 141	20.4	48.6	708	3	US-09-974-300-5134	Sequence 5134, Ap	C 214	20.2	47.6	695	12	US-10-301-480-564332	Sequence 564332,
C 142	20.4	48.6	939	11	US-10-932-182A-2422	Sequence 2422, Ap	C 215	20.2	47.6	995	12	US-10-301-480-117741	Sequence 117741,
C 143	20.4	48.6	1800	10	US-10-750-185-38727	Sequence 38727, A	C 216	20.2	47.6	1123	7	US-10-264-237-334	Sequence 334, App
C 144	20.4	48.6	1800	10	US-10-750-623-38727	Sequence 38727, A	C 217	20.2	47.6	1426	8	US-10-437-963-14884	Sequence 14884, A
C 145	20.4	48.6	1812	4	US-09-925-065A-700542	Sequence 700542, A	C 218	20.2	47.6	7890	7	US-10-311-455-1096	Sequence 1096, Ap
C 146	20.4	48.6	1812	5	US-09-925-065A-700542	Sequence 700542, A	C 219	20.2	47.6	10591	8	US-10-601-807-1	Sequence 1, Appl
C 147	20.4	48.6	2523	7	US-10-094-749-1277	Sequence 1277, Ap	C 220	20.2	47.6	38703	15	US-11-052-544-28	Sequence 28, Appl
C 148	20.4	48.6	2648	7	US-10-104-047-349	Sequence 349, App	221	19.8	47.1	176	8	US-10-242-535A-42990	Sequence 42990, A
C 149	20.4	48.6	2648	16	US-11-072-512-349	Sequence 349, App	222	19.8	47.1	176	8	US-10-085-783A-42990	Sequence 42990, A
C 150	20.4	48.6	2737	13	US-11-097-143-34528	Sequence 34528, A	223	19.8	47.1	524	4	US-09-925-065A-139905	Sequence 139905,
C 151	20.4	48.6	2781	10	US-10-956-157-4775	Sequence 4775, Ap	224	19.8	47.1	524	5	US-09-925-065A-139905	Sequence 139905,
C 152	20.4	48.6	2998	10	US-10-956-157-2547	Sequence 2547, Ap	C 225	19.8	47.1	536	12	US-10-301-480-234951	Sequence 234951,
C 153	20.4	48.6	3071	10	US-10-956-157-4776	Sequence 4776, Ap	C 226	19.8	47.1	536	12	US-10-301-480-848360	Sequence 848360,
C 154	20.4	48.6	122614	6	US-10-087-192-1726	Sequence 1726, Ap	227	19.8	47.1	542	4	US-09-925-065A-579267	Sequence 579267,
C 155	20.2	48.1	276	9	US-10-425-115-135315	Sequence 135315, A	228	19.8	47.1	542	5	US-09-925-065A-579267	Sequence 579267,
C 156	20.2	48.1	384	7	US-10-092-947A-7	Sequence 7, Appl	C 229	19.8	47.1	546	9	US-10-653-047-647	Sequence 647, App
C 157	20.2	48.1	537	12	US-10-301-480-336942	Sequence 336942, A	230	19.8	47.1	552	4	US-09-925-065A-770573	Sequence 770573,
C 158	20.2	48.1	537	12	US-10-301-480-336943	Sequence 336943, A	231	19.8	47.1	552	5	US-09-925-065A-770573	Sequence 770573,
C 159	20.2	48.1	537	12	US-10-301-480-336944	Sequence 336944, A	C 232	19.8	47.1	556	4	US-09-925-065A-773105	Sequence 773105,
C 160	20.2	48.1	537	12	US-10-301-480-950351	Sequence 950351, A	C 233	19.8	47.1	556	5	US-09-925-065A-773105	Sequence 773105,
C 161	20.2	48.1	537	12	US-10-301-480-950352	Sequence 950352, A	C 234	19.8	47.1	608	12	US-10-301-480-329656	Sequence 329656,
C 162	20.2	48.1	537	12	US-10-301-480-950353	Sequence 950353, A	C 235	19.8	47.1	608	12	US-10-301-480-943065	Sequence 943065,
C 163	20.2	48.1	542	4	US-09-925-065A-579266	Sequence 579266, A	C 236	19.8	47.1	609	4	US-09-925-065A-249821	Sequence 249821,

C 383	19.4	46.2	1725	3	US-09-946-290-7	Sequence 7, Appli	C 456	19.2	45.7	933	7	US-10-027-632-31762	Sequence 31762, A
C 384	19.4	46.2	1852	16	US-11-136-527-3981	Sequence 3981, Ap	457	19.2	45.7	959	3	US-09-899-046-49	Sequence 49, Appl
C 385	19.4	46.2	2173	13	US-11-097-143-7285	Sequence 7285, Ap	458	19.2	45.7	959	3	US-09-899-046-51	Sequence 51, Appl
C 386	19.4	46.2	2228	8	US-10-424-599-24936	Sequence 24936, A	459	19.2	45.7	959	3	US-09-899-046-53	Sequence 53, Appl
C 387	19.4	46.2	2425	10	US-10-450-763-15164	Sequence 15164, A	460	19.2	45.7	959	3	US-09-878-281-49	Sequence 49, Appl
C 388	19.4	46.2	2459	10	US-10-496-703-15164	Sequence 15164, A	461	19.2	45.7	959	3	US-09-878-281-51	Sequence 51, Appl
C 389	19.4	46.2	3142	13	US-11-097-143-19504	Sequence 19504, A	462	19.2	45.7	959	3	US-09-878-281-53	Sequence 53, Appl
C 390	19.4	46.2	3750	16	US-11-126-022-24	Sequence 24, Appl	463	19.2	45.7	959	3	US-09-873-224-49	Sequence 49, Appl
C 391	19.4	46.2	3895	6	US-10-087-192-245	Sequence 245, Appl	464	19.2	45.7	959	3	US-09-873-224-51	Sequence 51, Appl
C 392	19.4	46.2	6878	13	US-11-097-143-23545	Sequence 23545, A	465	19.2	45.7	959	3	US-09-873-224-53	Sequence 53, Appl
C 393	19.4	46.2	20210	3	US-09-764-870-598	Sequence 598, App	466	19.2	45.7	998	10	US-10-750-185-34930	Sequence 34930, A
C 394	19.4	46.2	20210	6	US-10-125-540-598	Sequence 598, App	467	19.2	45.7	998	10	US-10-750-623-34930	Sequence 34930, A
C 395	19.4	46.2	32591	7	US-10-085-117-187	Sequence 187, App	468	19.2	45.7	1000	12	US-10-301-480-572494	Sequence 572494, A
C 396	19.4	46.2	83836	6	US-10-087-192-1849	Sequence 1849, Ap	469	19.2	45.7	1000	12	US-10-301-480-1185903	Sequence 1185903, A
C 397	19.4	46.2	187090	8	US-10-672-76434-34	Sequence 34, Appl	470	19.2	45.7	1103	10	US-10-750-185-57665	Sequence 57665, A
C 398	19.4	46.2	187844	9	US-10-719-993-6883	Sequence 6883, Ap	471	19.2	45.7	1103	10	US-10-750-623-57665	Sequence 57665, A
C 399	19.4	46.2	196686	6	US-10-087-192-484	Sequence 484, App	472	19.2	45.7	1111	8	US-10-169-103-5	Sequence 5, Appli
C 400	19.4	46.2	198522	6	US-10-087-192-244	Sequence 244, App	473	19.2	45.7	1311	9	US-10-767-795-3808	Sequence 3808, Ap
C 401	19.4	46.2	495635	10	US-10-737-082-12	Sequence 12, Appl	474	19.2	45.7	1400	8	US-10-424-599-125443	Sequence 125443, A
C 402	19.4	46.2	495635	10	US-10-765-790-12	Sequence 12, Appl	475	19.2	45.7	1641	6	US-10-128-714-2384	Sequence 2384, Ap
C 403	19.4	46.2	705636	10	US-10-737-082-30	Sequence 30, Appl	476	19.2	45.7	1725	8	US-10-389-647-105	Sequence 105, App
C 404	19.4	46.2	705636	10	US-10-765-790-30	Sequence 30, Appl	477	19.2	45.7	2006	6	US-10-128-714-1384	Sequence 1384, Ap
C 405	19.2	45.7	252	9	US-10-357-930-7212	Sequence 7212, Ap	478	19.2	45.7	2070	6	US-10-128-714-1384	Sequence 7384, Ap
C 406	19.2	45.7	398	9	US-10-357-930-72180	Sequence 37180, A	479	19.2	45.7	2162	4	US-09-925-065A-674877	Sequence 674877, A
C 407	19.2	45.7	405	3	US-09-864-761-5411	Sequence 5411, Ap	480	19.2	45.7	2162	4	US-09-925-065A-674877	Sequence 674877, A
C 408	19.2	45.7	465	3	US-09-925-065A-555136	Sequence 555136, A	481	19.2	45.7	2162	5	US-09-925-065A-674878	Sequence 674878, A
C 409	19.2	45.7	504	4	US-09-925-065A-555136	Sequence 555136, A	482	19.2	45.7	2162	5	US-09-925-065A-674878	Sequence 674878, A
C 410	19.2	45.7	504	4	US-09-925-065A-555136	Sequence 555136, A	483	19.2	45.7	2162	5	US-09-925-065A-674878	Sequence 674878, A
C 411	19.2	45.7	507	9	US-10-767-795-3810	Sequence 3810, Ap	484	19.2	45.7	2330	8	US-10-424-599-124753	Sequence 124753, A
C 412	19.2	45.7	526	4	US-09-925-065A-654017	Sequence 654017, A	485	19.2	45.7	2565	10	US-10-750-185-50905	Sequence 50905, A
C 413	19.2	45.7	526	5	US-09-925-065A-654017	Sequence 654017, A	486	19.2	45.7	2565	10	US-10-750-623-50905	Sequence 50905, A
C 414	19.2	45.7	530	9	US-10-425-115-176062	Sequence 176062, A	487	19.2	45.7	3144	7	US-10-108-260A-1860	Sequence 1860, Ap
C 415	19.2	45.7	540	4	US-09-925-065A-125584	Sequence 125584, A	488	19.2	45.7	3257	10	US-10-974-440-103	Sequence 103, App
C 416	19.2	45.7	540	5	US-09-925-065A-125584	Sequence 125584, A	489	19.2	45.7	3507	10	US-10-887-583A-803	Sequence 803, App
C 417	19.2	45.7	548	6	US-10-027-632-282961	Sequence 282961, A	490	19.2	45.7	3675	13	US-11-097-143-19709	Sequence 19709, A
C 418	19.2	45.7	548	7	US-10-027-632-282961	Sequence 282961, A	491	19.2	45.7	3818	11	US-10-893-483-146	Sequence 146, App
C 419	19.2	45.7	548	12	US-10-301-480-223046	Sequence 223046, A	492	19.2	45.7	4005	6	US-10-128-714-384	Sequence 384, App
C 420	19.2	45.7	548	12	US-10-301-480-223046	Sequence 223046, A	493	19.2	45.7	4235	6	US-10-128-714-384	Sequence 384, Ap
C 421	19.2	45.7	558	4	US-09-925-065A-64077	Sequence 64077, A	494	19.2	45.7	4309	10	US-10-930-723A-17	Sequence 17, Appl
C 422	19.2	45.7	558	5	US-09-925-065A-64077	Sequence 64077, A	495	19.2	45.7	4309	10	US-10-927-951A-17	Sequence 17, Appl
C 423	19.2	45.7	558	12	US-10-301-480-165315	Sequence 165315, A	496	19.2	45.7	5191	6	US-10-171-581-168	Sequence 168, App
C 424	19.2	45.7	558	12	US-10-301-480-165315	Sequence 165315, A	497	19.2	45.7	5191	7	US-10-393-892-5	Sequence 5, Appli
C 425	19.2	45.7	578	8	US-10-021-323-3976	Sequence 3976, Ap	498	19.2	45.7	5191	7	US-10-394-382-5	Sequence 5, Appli
C 426	19.2	45.7	579	3	US-09-899-046-153	Sequence 153, App	499	19.2	45.7	5490	8	US-10-115-635-34	Sequence 34, Appl
C 427	19.2	45.7	579	3	US-09-899-046-155	Sequence 155, App	500	19.2	45.7	5981	6	US-10-044-090-291	Sequence 291, App
C 428	19.2	45.7	579	3	US-09-899-046-155	Sequence 155, App	501	19.2	45.7	5981	6	US-10-044-090-291	Sequence 291, App
C 429	19.2	45.7	579	3	US-09-899-046-155	Sequence 155, App	502	19.2	45.7	7619	13	US-11-097-143-19708	Sequence 19708, A
C 430	19.2	45.7	579	3	US-09-878-281-153	Sequence 153, App	503	19.2	45.7	9027	7	US-10-117-960-11	Sequence 11, Appl
C 431	19.2	45.7	579	3	US-09-878-281-155	Sequence 155, App	504	19.2	45.7	9209	7	US-10-117-960-14	Sequence 14, Appl
C 432	19.2	45.7	579	3	US-09-878-281-195	Sequence 195, App	505	19.2	45.7	10551	7	US-10-117-960-1	Sequence 1, Appli
C 433	19.2	45.7	579	3	US-09-873-224-153	Sequence 153, App	506	19.2	45.7	54333	7	US-10-085-117-124	Sequence 124, App
C 434	19.2	45.7	579	3	US-09-873-224-155	Sequence 155, App	507	19.2	45.7	55689	8	US-10-741-601-5684	Sequence 5684, Ap
C 435	19.2	45.7	579	3	US-09-873-224-195	Sequence 195, App	508	19.2	45.7	55689	9	US-10-741-600-17734	Sequence 17734, A
C 436	19.2	45.7	580	3	US-09-899-046-45	Sequence 45, Appl	509	19.2	45.7	55689	10	US-10-995-561-13318	Sequence 13318, A
C 437	19.2	45.7	580	3	US-09-899-046-47	Sequence 47, Appl	510	19.2	45.7	72691	9	US-10-741-600-17712	Sequence 17712, A
C 438	19.2	45.7	580	3	US-09-878-281-45	Sequence 45, Appl	511	19.2	45.7	106323	6	US-09-803-661-3	Sequence 3, Appli
C 439	19.2	45.7	580	3	US-09-878-281-47	Sequence 47, Appl	512	19.2	45.7	106323	6	US-10-300-827-3	Sequence 3, Appli
C 440	19.2	45.7	580	3	US-09-873-224-45	Sequence 45, Appl	513	19.2	45.7	155083	11	US-10-330-773-868	Sequence 868, App
C 441	19.2	45.7	580	6	US-10-027-632-106142	Sequence 106142, A	514	19.2	45.7	169725	15	US-11-121-086-63	Sequence 63, Appl
C 442	19.2	45.7	580	6	US-10-027-632-106142	Sequence 106142, A	515	19.2	45.7	200628	15	US-11-121-086-62	Sequence 62, Appl
C 443	19.2	45.7	590	4	US-09-925-065A-620406	Sequence 620406, A	516	19.2	45.7	202251	6	US-10-087-192-985	Sequence 985, App
C 444	19.2	45.7	590	5	US-09-925-065A-620406	Sequence 620406, A	517	19.2	45.7	202251	6	US-10-087-192-985	Sequence 985, App
C 445	19.2	45.7	630	3	US-09-873-025-31	Sequence 31, Appl	518	19.2	45.7	19	8	US-10-600-816-32	Sequence 32, Appl
C 446	19.2	45.7	630	3	US-09-899-303-31	Sequence 31, Appl	519	19.2	45.7	19	14	US-11-083-784-618880	Sequence 618880, A
C 447	19.2	45.7	630	3	US-09-995-808-31	Sequence 31, Appl	520	19.2	45.7	19	14	US-11-083-784-618951	Sequence 618951, A
C 448	19.2	45.7	630	3	US-09-995-808-31	Sequence 31, Appl	521	19.2	45.7	19	15	US-11-101-244-618880	Sequence 618880, A
C 449	19.2	45.7	630	3	US-09-995-791-31	Sequence 31, Appl	522	19.2	45.7	19	15	US-11-101-244-618935	Sequence 618935, A
C 450	19.2	45.7	630	8	US-10-321-799-31	Sequence 31, Appl	523	19.2	45.7	19	15	US-11-101-244-618951	Sequence 618951, A
C 451	19.2	45.7	647	4	US-09-925-065A-907883	Sequence 907883, A	524	19.2	45.7	157	8	US-10-424-599-40918	Sequence 40918, A
C 452	19.2	45.7	647	5	US-09-925-065A-907883	Sequence 907883, A	525	19.2	45.7	201	10	US-10-995-561-67707	Sequence 67707, A
C 453	19.2	45.7	678	6	US-10-027-632-20810	Sequence 20810, A	526	19.2	45.7	242	7	US-10-026-925-52	Sequence 52, Appl
C 454	19.2	45.7	678	7	US-10-027-632-20810	Sequence 20810, A	527	19.2	45.7	289	3	US-09-294-093B-552	Sequence 552, App
C 455	19.2	45.7	933	6	US-10-027-632-31762	Sequence 31762, A	528	19.2	45.7				

C 529	19	45.2	456	3	US-09-918-995-27565	Sequence 27565, A	C 602	18.8	44.8	239	3	US-09-437-458-17	Sequence 17, Appl
C 530	19	45.2	467	9	US-10-425-115-65545	Sequence 65545, A	C 603	18.8	44.8	239	10	US-10-963-062-17	Sequence 17, Appl
C 531	19	45.2	482	3	US-09-918-995-29074	Sequence 29074, A	C 604	18.8	44.8	279	7	US-10-029-386-23952	Sequence 23952, A
C 532	19	45.2	489	3	US-09-930-213-147	Sequence 147, App	C 605	18.8	44.8	367	9	US-10-425-115-17325	Sequence 17325, A
C 533	19	45.2	489	3	US-09-930-213-323	Sequence 323, App	C 606	18.8	44.8	393	6	US-10-083-357-16	Sequence 16, Appl
C 534	19	45.2	540	4	US-09-928-065A-408940	Sequence 408940, A	C 607	18.8	44.8	466	3	US-09-960-352-6662	Sequence 6662, Ap
C 535	19	45.2	549	12	US-09-925-065A-408940	Sequence 408940, A	C 608	18.8	44.8	508	7	US-10-029-386-10235	Sequence 10235, A
C 536	19	45.2	549	12	US-10-301-480-23308	Sequence 23308, A	C 609	18.8	44.8	512	4	US-09-925-065A-163017	Sequence 163017, A
C 537	19	45.2	549	12	US-10-301-480-475435	Sequence 475435, A	C 610	18.8	44.8	512	5	US-09-925-065A-163017	Sequence 163017, A
C 538	19	45.2	549	12	US-10-301-480-636717	Sequence 636717, A	C 611	18.8	44.8	531	8	US-10-424-599-88340	Sequence 88340, A
C 539	19	45.2	549	12	US-10-301-480-1088844	Sequence 1088844, A	C 612	18.8	44.8	550	12	US-10-301-480-255534	Sequence 255534, A
C 540	19	45.2	551	12	US-10-301-480-464184	Sequence 464184, A	C 613	18.8	44.8	550	12	US-10-301-480-868943	Sequence 868943, A
C 541	19	45.2	551	12	US-10-301-480-464185	Sequence 464185, A	C 614	18.8	44.8	593	4	US-09-925-065A-667322	Sequence 667322, A
C 542	19	45.2	551	12	US-10-301-480-1077593	Sequence 1077593, A	C 615	18.8	44.8	593	5	US-09-925-065A-667322	Sequence 667322, A
C 543	19	45.2	551	12	US-10-301-480-1077594	Sequence 1077594, A	C 616	18.8	44.8	596	4	US-09-925-065A-741533	Sequence 741533, A
C 544	19	45.2	554	4	US-09-925-065A-396412	Sequence 396412, A	C 617	18.8	44.8	596	5	US-09-925-065A-741533	Sequence 741533, A
C 545	19	45.2	554	4	US-09-925-065A-396413	Sequence 396413, A	C 618	18.8	44.8	597	3	US-09-864-761-6862	Sequence 6862, Ap
C 546	19	45.2	554	5	US-09-925-065A-396412	Sequence 396412, A	C 619	18.8	44.8	597	4	US-09-925-065A-17936	Sequence 17936, A
C 547	19	45.2	554	5	US-09-925-065A-396413	Sequence 396413, A	C 620	18.8	44.8	597	5	US-09-925-065A-17936	Sequence 17936, A
C 548	19	45.2	600	10	US-10-972-079-84104	Sequence 84104, A	C 621	18.8	44.8	597	12	US-10-301-480-119173	Sequence 119173, A
C 549	19	45.2	600	10	US-10-750-185-3382	Sequence 3382, Ap	C 622	18.8	44.8	597	12	US-10-301-480-733582	Sequence 733582, A
C 550	19	45.2	600	10	US-10-750-623-3382	Sequence 3382, Ap	C 623	18.8	44.8	604	6	US-10-027-632-264086	Sequence 264086, A
C 551	19	45.2	617	4	US-09-925-065A-312371	Sequence 312371, A	C 624	18.8	44.8	604	6	US-10-027-632-264087	Sequence 264087, A
C 552	19	45.2	617	5	US-09-925-065A-312371	Sequence 312371, A	C 625	18.8	44.8	604	6	US-10-027-632-264088	Sequence 264088, A
C 553	19	45.2	617	12	US-10-301-480-387614	Sequence 387614, A	C 626	18.8	44.8	604	7	US-10-027-632-264088	Sequence 264088, A
C 554	19	45.2	617	12	US-10-301-480-1001023	Sequence 1001023, A	C 627	18.8	44.8	604	7	US-10-027-632-264087	Sequence 264087, A
C 555	19	45.2	742	12	US-10-301-480-610759	Sequence 610759, A	C 628	18.8	44.8	604	7	US-10-027-632-264088	Sequence 264088, A
C 556	19	45.2	742	12	US-10-301-480-1224168	Sequence 1224168, A	C 629	18.8	44.8	604	12	US-10-301-480-98910	Sequence 98910, A
C 557	19	45.2	752	4	US-09-925-065A-949141	Sequence 949141, A	C 630	18.8	44.8	604	12	US-10-301-480-98911	Sequence 98911, A
C 558	19	45.2	752	5	US-09-925-065A-949141	Sequence 949141, A	C 631	18.8	44.8	604	12	US-10-301-480-98912	Sequence 98912, A
C 559	19	45.2	763	5	US-09-925-065A-71093	Sequence 71093, A	C 632	18.8	44.8	604	12	US-10-301-480-712319	Sequence 712319, A
C 560	19	45.2	763	5	US-09-925-065A-71093	Sequence 71093, A	C 633	18.8	44.8	604	12	US-10-301-480-712320	Sequence 712320, A
C 561	19	45.2	763	12	US-10-301-480-172332	Sequence 172332, A	C 634	18.8	44.8	604	12	US-10-301-480-712321	Sequence 712321, A
C 562	19	45.2	763	12	US-10-301-480-785741	Sequence 785741, A	C 635	18.8	44.8	608	4	US-09-925-065A-236678	Sequence 236678, A
C 563	19	45.2	890	6	US-10-027-632-137471	Sequence 137471, A	C 636	18.8	44.8	608	5	US-09-925-065A-236678	Sequence 236678, A
C 564	19	45.2	1224	7	US-10-767-795-4226	Sequence 4226, Ap	C 637	18.8	44.8	610	4	US-09-925-065A-601174	Sequence 601174, A
C 565	19	45.2	1224	9	US-10-767-795-4226	Sequence 4226, Ap	C 638	18.8	44.8	610	5	US-09-925-065A-601174	Sequence 601174, A
C 566	19	45.2	1372	9	US-10-357-819-21	Sequence 21, Appl	C 639	18.8	44.8	614	12	US-10-301-480-932657	Sequence 932657, A
C 567	19	45.2	1664	7	US-10-305-720-1220	Sequence 1220, Ap	C 640	18.8	44.8	614	12	US-10-301-480-932666	Sequence 932666, A
C 568	19	45.2	1910	9	US-10-425-115-36196	Sequence 36196, A	C 641	18.8	44.8	624	6	US-10-027-632-192110	Sequence 192110, A
C 569	19	45.2	1998	3	US-09-800-629A-195	Sequence 195, App	C 642	18.8	44.8	624	6	US-10-027-632-192111	Sequence 192111, A
C 570	19	45.2	1998	8	US-10-679-532-195	Sequence 195, App	C 643	18.8	44.8	624	6	US-10-027-632-192111	Sequence 192111, A
C 571	19	45.2	2006	3	US-09-800-629A-176	Sequence 176, App	C 644	18.8	44.8	628	4	US-09-925-065A-192111	Sequence 192111, A
C 572	19	45.2	2006	8	US-10-679-532-176	Sequence 176, App	C 645	18.8	44.8	628	4	US-09-925-065A-664839	Sequence 664839, A
C 573	19	45.2	2024	8	US-10-641-643-1315	Sequence 1315, Ap	C 646	18.8	44.8	628	5	US-09-925-065A-664839	Sequence 664839, A
C 574	19	45.2	2205	13	US-11-097-143-4361	Sequence 4361, Ap	C 647	18.8	44.8	633	4	US-09-925-065A-796860	Sequence 796860, A
C 575	19	45.2	2212	9	US-10-478-146-20	Sequence 20, Appl	C 648	18.8	44.8	633	5	US-09-925-065A-796860	Sequence 796860, A
C 576	19	45.2	2297	3	US-09-853-753-1	Sequence 1, Appl	C 649	18.8	44.8	633	9	US-10-363-345A-29395	Sequence 29395, A
C 577	19	45.2	2397	7	US-10-104-047-659	Sequence 659, App	C 650	18.8	44.8	633	10	US-10-363-345A-29395	Sequence 29395, A
C 578	19	45.2	2397	16	US-11-072-512-659	Sequence 659, App	C 651	18.8	44.8	633	10	US-10-363-345A-29395	Sequence 29395, A
C 579	19	45.2	2547	10	US-10-450-763-11192	Sequence 11192, A	C 652	18.8	44.8	633	10	US-10-363-345A-29395	Sequence 29395, A
C 580	19	45.2	2798	8	US-10-437-963-65385	Sequence 65385, A	C 653	18.8	44.8	638	4	US-09-925-065A-850798	Sequence 850798, A
C 581	19	45.2	2876	7	US-10-108-260A-1526	Sequence 1526, Ap	C 654	18.8	44.8	638	5	US-09-925-065A-850798	Sequence 850798, A
C 582	19	45.2	3106	10	US-10-750-185-53364	Sequence 53364, A	C 655	18.8	44.8	677	10	US-10-750-185-47137	Sequence 47137, A
C 583	19	45.2	3106	10	US-10-750-623-53364	Sequence 53364, A	C 656	18.8	44.8	677	10	US-10-750-623-47137	Sequence 47137, A
C 584	19	45.2	3345	13	US-11-097-143-4360	Sequence 4360, Ap	C 657	18.8	44.8	746	6	US-10-027-632-103932	Sequence 103932, A
C 585	19	45.2	4305	9	US-10-741-600-17924	Sequence 17924, A	C 658	18.8	44.8	746	6	US-10-027-632-103932	Sequence 103932, A
C 586	19	45.2	15373	11	US-10-330-773-146	Sequence 146, App	C 659	18.8	44.8	746	6	US-10-027-632-103932	Sequence 103932, A
C 587	19	45.2	45555	9	US-10-741-600-17798	Sequence 17798, A	C 660	18.8	44.8	746	6	US-10-027-632-323704	Sequence 323704, A
C 588	19	45.2	51138	11	US-10-741-600-18004	Sequence 18004, A	C 661	18.8	44.8	746	7	US-10-027-632-103932	Sequence 103932, A
C 589	19	45.2	101169	11	US-10-330-773-561	Sequence 561, App	C 662	18.8	44.8	746	7	US-10-027-632-103932	Sequence 103932, A
C 590	19	45.2	210920	11	US-10-330-773-99	Sequence 99, Appl	C 663	18.8	44.8	746	7	US-10-027-632-323675	Sequence 323675, A
C 591	19	45.2	398287	8	US-10-741-601-5719	Sequence 5719, Ap	C 664	18.8	44.8	746	7	US-10-027-632-323704	Sequence 323704, A
C 592	19	45.2	398287	9	US-10-741-600-17839	Sequence 17839, A	C 665	18.8	44.8	750	8	US-10-335-977-426	Sequence 426, App
C 593	19	45.2	398287	10	US-10-995-561-13396	Sequence 13396, A	C 666	18.8	44.8	750	8	US-10-335-977-427	Sequence 427, App
C 594	18.8	44.8	25	10	US-10-956-157-254203	Sequence 254203, A	C 667	18.8	44.8	755	10	US-10-750-185-47948	Sequence 47948, A
C 595	18.8	44.8	25	10	US-10-956-157-254203	Sequence 254203, A	C 668	18.8	44.8	755	10	US-10-750-623-47948	Sequence 47948, A
C 596	18.8	44.8	72	11	US-10-956-157-278010	Sequence 278010, A	C 669	18.8	44.8	777	4	US-09-925-065A-78367	Sequence 78367, A
C 597	18.8	44.8	80	8	US-10-384-245-5	Sequence 245, Ap	C 670	18.8	44.8	777	5	US-09-925-065A-78367	Sequence 78367, A
C 598	18.8	44.8	131	3	US-09-864-761-23595	Sequence 23595, A	C 671	18.8	44.8	777	12	US-10-301-480-179606	Sequence 179606, A
C 599	18.8	44.8	201	16	US-11-124-3678A-19437	Sequence 19437, A	C 672	18.8	44.8	777	12	US-10-301-480-793015	Sequence 793015, A
C 600	18.8	44.8	220	8	US-10-424-599-102922	Sequence 102922, A	C 673	18.8	44.8	783	8	US-10-335-977-428	Sequence 428, App
C 601	18.8	44.8	239	3	US-09-729-447-17	Sequence 17, Appl	C 674	18.8	44.8	821	4	US-09-925-065A-67322	Sequence 67322, A

821	18.6	44.3	489	4	US-09-925-065A-907169	Sequence 907169,	894	18.6	44.3	645	12	US-10-301-480-42939	Sequence 42939, A
822	18.6	44.3	489	4	US-09-925-065A-907170	Sequence 907170,	895	18.6	44.3	645	12	US-10-301-480-656348	Sequence 656348,
823	18.6	44.3	489	4	US-09-925-065A-907169	Sequence 907169,	896	18.6	44.3	653	7	US-10-369-493-34922	Sequence 34922, A
824	18.6	44.3	489	5	US-09-925-065A-907170	Sequence 907170,	897	18.6	44.3	656	9	US-10-369-493-3684	Sequence 3684, Ap
825	18.6	44.3	491	10	US-10-972-079-2032	Sequence 2032, Ap	898	18.6	44.3	660	4	US-09-925-065A-76623	Sequence 76623, A
826	18.6	44.3	500	3	US-09-918-995-3085	Sequence 3085, Ap	899	18.6	44.3	660	4	US-09-925-065A-76624	Sequence 76624, A
827	18.6	44.3	501	6	US-10-027-632-276199	Sequence 276199,	900	18.6	44.3	660	4	US-09-925-065A-76625	Sequence 76625, A
828	18.6	44.3	501	7	US-10-027-632-276199	Sequence 276199,	901	18.6	44.3	660	4	US-09-925-065A-76626	Sequence 76626, A
829	18.6	44.3	528	4	US-09-925-065A-554502	Sequence 554502,	902	18.6	44.3	660	5	US-09-925-065A-76623	Sequence 76623, A
830	18.6	44.3	528	4	US-09-925-065A-554503	Sequence 554503,	903	18.6	44.3	660	5	US-09-925-065A-76624	Sequence 76624, A
831	18.6	44.3	528	4	US-09-925-065A-554504	Sequence 554504,	904	18.6	44.3	660	5	US-09-925-065A-76625	Sequence 76625, A
832	18.6	44.3	528	4	US-09-925-065A-554505	Sequence 554505,	905	18.6	44.3	660	5	US-09-925-065A-76626	Sequence 76626, A
833	18.6	44.3	528	5	US-09-925-065A-554502	Sequence 554502,	906	18.6	44.3	660	12	US-10-301-480-177862	Sequence 177862,
834	18.6	44.3	528	5	US-09-925-065A-554503	Sequence 554503,	907	18.6	44.3	660	12	US-10-301-480-177863	Sequence 177863,
835	18.6	44.3	528	5	US-09-925-065A-554504	Sequence 554504,	908	18.6	44.3	660	12	US-10-301-480-177864	Sequence 177864,
836	18.6	44.3	528	5	US-09-925-065A-554505	Sequence 554505,	909	18.6	44.3	660	12	US-10-301-480-177865	Sequence 177865,
837	18.6	44.3	533	8	US-10-021-323-10817	Sequence 10817, A	910	18.6	44.3	660	12	US-10-301-480-791271	Sequence 791271,
838	18.6	44.3	538	4	US-09-925-065A-703626	Sequence 703626,	911	18.6	44.3	660	12	US-10-301-480-791272	Sequence 791272,
839	18.6	44.3	538	4	US-09-925-065A-703627	Sequence 703627,	912	18.6	44.3	660	12	US-10-301-480-791273	Sequence 791273,
840	18.6	44.3	538	4	US-09-925-065A-703628	Sequence 703628,	913	18.6	44.3	660	12	US-10-301-480-791274	Sequence 791274,
841	18.6	44.3	538	5	US-09-925-065A-703626	Sequence 703626,	914	18.6	44.3	662	6	US-10-027-632-181867	Sequence 181867,
842	18.6	44.3	538	5	US-09-925-065A-703627	Sequence 703627,	915	18.6	44.3	662	6	US-10-027-632-181868	Sequence 181868,
843	18.6	44.3	538	5	US-09-925-065A-703628	Sequence 703628,	916	18.6	44.3	662	7	US-10-027-632-181868	Sequence 181868,
844	18.6	44.3	565	12	US-10-301-480-522960	Sequence 522960,	917	18.6	44.3	662	7	US-10-027-632-181868	Sequence 181868,
845	18.6	44.3	565	12	US-10-301-480-1136369	Sequence 1136369,	918	18.6	44.3	666	4	US-09-925-065A-596206	Sequence 596206,
846	18.6	44.3	570	4	US-09-925-065A-467403	Sequence 467403,	919	18.6	44.3	666	4	US-09-925-065A-596214	Sequence 596214,
847	18.6	44.3	570	4	US-09-925-065A-467403	Sequence 467403,	920	18.6	44.3	666	4	US-09-925-065A-596215	Sequence 596215,
848	18.6	44.3	571	10	US-10-972-079-2031	Sequence 2031, Ap	921	18.6	44.3	666	4	US-09-925-065A-596216	Sequence 596216,
849	18.6	44.3	579	3	US-09-899-046-179	Sequence 179, App	922	18.6	44.3	666	5	US-09-925-065A-596206	Sequence 596206,
850	18.6	44.3	579	3	US-09-878-281-179	Sequence 179, App	923	18.6	44.3	666	5	US-09-925-065A-596214	Sequence 596214,
851	18.6	44.3	579	3	US-09-873-224-179	Sequence 179, App	924	18.6	44.3	666	5	US-09-925-065A-596215	Sequence 596215,
852	18.6	44.3	579	12	US-10-301-480-458453	Sequence 458453,	925	18.6	44.3	666	5	US-09-925-065A-596216	Sequence 596216,
853	18.6	44.3	579	12	US-10-301-480-1071862	Sequence 1071862,	926	18.6	44.3	674	4	US-09-925-065A-823799	Sequence 823799,
854	18.6	44.3	580	4	US-09-925-065A-390127	Sequence 390127,	927	18.6	44.3	674	5	US-09-925-065A-823799	Sequence 823799,
855	18.6	44.3	580	5	US-09-925-065A-390127	Sequence 390127,	928	18.6	44.3	741	9	US-10-767-795-4770	Sequence 4770, Ap
856	18.6	44.3	585	4	US-09-925-065A-565229	Sequence 565229,	929	18.6	44.3	832	6	US-10-198-846-9427	Sequence 9427, Ap
857	18.6	44.3	585	5	US-09-925-065A-565229	Sequence 565229,	930	18.6	44.3	904	8	US-10-424-599-9747	Sequence 9747, Ap
858	18.6	44.3	586	4	US-09-925-065A-305680	Sequence 305680,	931	18.6	44.3	943	7	US-10-027-632-164347	Sequence 164347,
859	18.6	44.3	586	4	US-09-925-065A-305681	Sequence 305681,	932	18.6	44.3	943	7	US-10-027-632-164347	Sequence 164347,
860	18.6	44.3	586	5	US-09-925-065A-305680	Sequence 305680,	933	18.6	44.3	977	12	US-10-301-480-557642	Sequence 557642,
861	18.6	44.3	586	5	US-09-925-065A-305681	Sequence 305681,	934	18.6	44.3	977	12	US-10-301-480-117051	Sequence 117051,
862	18.6	44.3	591	4	US-09-925-065A-261130	Sequence 261130,	935	18.6	44.3	984	16	US-11-136-527-21	Sequence 21, Appl
863	18.6	44.3	591	5	US-09-925-065A-261130	Sequence 261130,	936	18.6	44.3	998	12	US-10-301-480-534811	Sequence 534811,
864	18.6	44.3	592	4	US-09-925-065A-751227	Sequence 751227,	937	18.6	44.3	998	12	US-10-301-480-534812	Sequence 534812,
865	18.6	44.3	592	5	US-09-925-065A-751227	Sequence 751227,	938	18.6	44.3	998	12	US-10-301-480-1148220	Sequence 1148220,
866	18.6	44.3	594	12	US-10-301-480-381675	Sequence 381675,	939	18.6	44.3	998	12	US-10-301-480-1148221	Sequence 1148221,
867	18.6	44.3	594	12	US-10-301-480-381676	Sequence 381676,	940	18.6	44.3	1194	13	US-11-097-143-18002	Sequence 18002, A
868	18.6	44.3	594	12	US-10-301-480-995084	Sequence 995084,	941	18.6	44.3	1371	8	US-10-282-122A-42265	Sequence 42265, A
869	18.6	44.3	594	12	US-10-301-480-995085	Sequence 995085,	942	18.6	44.3	1371	8	US-10-437-963-53068	Sequence 53068, A
870	18.6	44.3	597	4	US-09-925-065A-105081	Sequence 105081,	943	18.6	44.3	1494	3	US-09-764-891-5888	Sequence 5888, Ap
871	18.6	44.3	597	4	US-09-925-065A-105081	Sequence 105081,	944	18.6	44.3	1496	3	US-09-764-891-5888	Sequence 5888, Ap
872	18.6	44.3	597	5	US-09-925-065A-105080	Sequence 105080,	945	18.6	44.3	1496	3	US-09-764-891-5888	Sequence 5888, Ap
873	18.6	44.3	597	5	US-09-925-065A-105081	Sequence 105081,	946	18.6	44.3	1496	3	US-09-764-891-5888	Sequence 5888, Ap
874	18.6	44.3	598	4	US-09-925-065A-759029	Sequence 759029,	947	18.6	44.3	1633	8	US-10-437-963-81484	Sequence 81484, A
875	18.6	44.3	598	5	US-09-925-065A-759029	Sequence 759029,	948	18.6	44.3	1642	11	US-10-522-096-5	Sequence 5, Appl1
876	18.6	44.3	598	12	US-10-301-480-339452	Sequence 339452,	949	18.6	44.3	1862	4	US-09-925-065A-681082	Sequence 681082,
877	18.6	44.3	598	12	US-10-301-480-398297	Sequence 398297,	950	18.6	44.3	1862	4	US-09-925-065A-681082	Sequence 681082,
878	18.6	44.3	598	12	US-10-301-480-952861	Sequence 952861,	951	18.6	44.3	1905	11	US-10-932-182A-1386	Sequence 1386, Ap
879	18.6	44.3	598	12	US-10-301-480-1011706	Sequence 1011706,	952	18.6	44.3	1908	7	US-10-369-493-37280	Sequence 37280, A
880	18.6	44.3	600	10	US-10-972-079-2029	Sequence 2029, Ap	953	18.6	44.3	2046	7	US-10-369-493-33381	Sequence 33381, A
881	18.6	44.3	600	10	US-10-972-079-2029	Sequence 2029, Ap	954	18.6	44.3	2046	7	US-10-369-493-33381	Sequence 33381, A
882	18.6	44.3	600	16	US-11-136-527-4117	Sequence 4117, Ap	955	18.6	44.3	2271	13	US-11-097-143-29806	Sequence 29806, A
883	18.6	44.3	601	12	US-10-301-480-205200	Sequence 205200,	956	18.6	44.3	2316	13	US-11-097-143-29806	Sequence 29806, A
884	18.6	44.3	601	12	US-10-301-480-205201	Sequence 205201,	957	18.6	44.3	2553	4	US-09-925-065A-699896	Sequence 699896,
885	18.6	44.3	601	12	US-10-301-480-818609	Sequence 818609,	958	18.6	44.3	2553	5	US-09-925-065A-699896	Sequence 699896,
886	18.6	44.3	601	12	US-10-301-480-818610	Sequence 818610,	959	18.6	44.3	2604	7	US-10-369-493-33688	Sequence 33688, A
887	18.6	44.3	602	6	US-10-027-632-256523	Sequence 256523,	960	18.6	44.3	2615	13	US-11-097-143-11888	Sequence 11888, A
888	18.6	44.3	602	7	US-10-027-632-256523	Sequence 256523,	961	18.6	44.3	2787	5	US-09-925-065A-554597	Sequence 554597,
889	18.6	44.3	602	12	US-10-301-480-91320	Sequence 91320, A	962	18.6	44.3	2787	5	US-09-925-065A-554597	Sequence 554597,
890	18.6	44.3	602	12	US-10-301-480-91320	Sequence 91320, A	963	18.6	44.3	3034	13	US-11-097-143-14149	Sequence 14149, A
891	18.6	44.3	603	3	US-09-974-300-798	Sequence 798, App	964	18.6	44.3	3306	13	US-11-097-143-18001	Sequence 18001, A
892	18.6	44.3	645	4	US-09-925-065A-838215	Sequence 838215,	965	18.6	44.3	3314	8	US-10-424-599-17418	Sequence 17418, A
893	18.6	44.3	645	5	US-09-925-065A-838215	Sequence 838215,	966	18.6	44.3	3478	16	US-11-024-959-64	Sequence 64, Appl1

```
c 967 18.6 44.3 4806 13 US-11-097-143-11887
968 18.6 44.3 5082 13 US-11-097-143-4180
969 18.6 44.3 6975 13 US-11-097-143-11656
970 18.6 44.3 10343 13 US-11-097-143-118610
c 971 18.6 44.3 24678 9 US-10-741-600-17702
c 972 18.6 44.3 50000 6 US-10-152-724A-22
973 18.6 44.3 96583 15 US-11-117-187-203
c 974 18.6 44.3 96595 7 US-10-034-650-34
975 18.6 44.3 117329 10 US-10-987-384-1
c 976 18.6 44.3 127369 6 US-10-087-192-238
977 18.6 44.3 180654 15 US-11-121-086-58
978 18.6 44.3 197241 16 US-11-114-798-47
979 18.6 44.3 215221 6 US-10-087-192-1360
980 18.6 44.3 218336 9 US-10-473-939-10
981 18.6 44.3 310122 9 US-10-417-375-1
c 982 18.6 44.3 382256 10 US-10-820-226-1
983 18.6 44.3 382259 13 US-11-029-984-1
c 984 18.6 44.3 559686 5 US-09-824-575A-3
c 985 18.6 44.3 653458 10 US-10-461-862-4
986 18.6 44.3 1223197 6 US-10-027-632-179264
987 18.6 44.3 1223197 7 US-10-027-632-179264
988 18.4 43.8 201 10 US-10-995-561-64681
c 989 18.4 43.8 300 10 US-10-779-543-3237
c 990 18.4 43.8 369 16 US-11-102-512-110
991 18.4 43.8 371 9 US-10-425-115-82370
992 18.4 43.8 444 12 US-10-301-480-9374
993 18.4 43.8 444 12 US-10-301-480-9375
994 18.4 43.8 444 12 US-10-301-480-622783
995 18.4 43.8 444 12 US-10-301-480-622784
996 18.4 43.8 445 4 US-09-925-065A-248267
997 18.4 43.8 445 5 US-09-925-065A-248267
c 998 18.4 43.8 448 3 US-09-960-352-8299
999 18.4 43.8 451 12 US-10-301-480-328272
1000 18.4 43.8 451 12 US-10-301-480-941681
```

ALIGNMENTS

```
RESULT 1
US-10-600-816-30
; Sequence 30, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-30
```

```
Query Match 100.0%; Score 42; DB 8; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
```

```
Db 1 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
```

RESULT 2

```
US-10-066-543-1937
; Sequence 1937, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1937
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1937

Query Match 100.0%; Score 42; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 7.8e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 344 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 385
|||||
```

RESULT 3

```
US-10-066-543-1811/c
; Sequence 1811, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1811
```

```
Query Match 100.0%; Score 42; DB 6; Length 552;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 154 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 113
|||||
```

RESULT 4

US-09-969-034-2222/c
; Sequence 2222, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2222
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 491, 500, 517, 530, 545, 547, 553, 558, 570, 599, 620
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-2222

Query Match 100.0%; Score 42; DB 3; Length 620;
Best Local Similarity 100.0%; Pred. No. 8.2e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 154 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 113

RESULT 5

US-09-969-034-3393/c
; Sequence 3393, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3393
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 363, 470, 472, 516, 523, 531, 551, 567, 573, 588, 594, 613,
; LOCATION: 623

; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3393

Query Match 100.0%; Score 42; DB 3; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 158 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 117

RESULT 6

US-10-125-968-701/c
; Sequence 701, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-032
; CURRENT APPLICATION NUMBER: US/10/125,968
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,163
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 701
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 72, 76, 77, 99, 101, 423, 445, 468, 489, 557, 566, 633
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-701

Query Match 100.0%; Score 42; DB 7; Length 642;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 184 GCCACGCTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 143

RESULT 7

US-10-712-615-134
; Sequence 134, Application US/10712615
; Publication No. US20040214317A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBM Y8, EXPRESSED
; FILE REFERENCE: D0047A-CIP
; CURRENT APPLICATION NUMBER: US/10/712,615
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: U.S. 09/992,238
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: U.S. 60/248,285
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: U.S. 60/268,581
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: U.S. 60/308,285
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: U.S. 60/317,166
; PRIOR FILING DATE: 2001-09-04

```
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 134
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: dna;Homo sapiens
; US-10-712-615-134

Query Match          100.0%; Score 42; DB 9; Length 1071;
Best Local Similarity 100.0%; Pred. No. 9.4e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 1018 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1059

RESULT 8
US-09-866-050A-249
; Sequence 249, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
; US-09-866-050A-249

Query Match          100.0%; Score 42; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 990 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 9
US-10-152-661-249
; Sequence 249, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
```

```
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
; US-10-152-661-249

Query Match          100.0%; Score 42; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 990 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 10
US-10-313-542-223
; Sequence 223, Application US/10313542
; Publication No. US20030120057A1
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/10/313,542
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/495,050
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030120057A1 2681738CT1
; US-10-313-542-223

Query Match          100.0%; Score 42; DB 7; Length 1228;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 541 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 582

RESULT 11
US-11-060-756-2418
; Sequence 2418, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
```



```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2418
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-2418

Query Match      100.0%; Score 42; DB 13; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 215 GCCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 256

RESULT 12
US-11-060-756-6690
; Sequence 6690, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6690
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-6690

Query Match      100.0%; Score 42; DB 13; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 215 GCCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 256

RESULT 13
US-10-936-626-64
; Sequence 64, Application US/10936626
; Publication No. US20050106644A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; CURRENT FILING DATE: 2004-06-21
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
```

```
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-64

Query Match      100.0%; Score 42; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
    |||||||
Db 1145 GCCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1186

RESULT 14
US-10-938-061-64
; Sequence 64, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie
; APPLICANT: Chuntharapal, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; FILE REFERENCE: P5001R1P1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
```

```
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-938-061-64

Query Match      100.0%; Score 42; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
   |||||||
Db 1145 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1186

RESULT 15
US-10-224-289-5
; Sequence 5, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-5

Query Match      100.0%; Score 42; DB 7; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
   |||||||
Db 1137 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1178

RESULT 16
US-10-935-190-43
; Sequence 43, Application US/10935190
; Publication No. US20050037466A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BURFORD, Neil
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyrang Aina M.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; TITLE OF INVENTION: RECEPTORS AND ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0726 PCT
; CURRENT APPLICATION NUMBER: US/10/935,190
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US/10/031,904
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/145,232; 60/158,578; 60/165,192
; PRIOR FILING DATE: 1999-07-21; 1999-10-07; 1999-11-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 43
```

```
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-935-190-43

Query Match      100.0%; Score 42; DB 9; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
   |||||||
Db 1137 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1178

RESULT 17
US-10-775-920-13
; Sequence 13, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-13

Query Match      100.0%; Score 42; DB 9; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
   |||||||
Db 1125 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1166

RESULT 18
US-10-224-289-3
; Sequence 3, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; TITLE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-3

Query Match      100.0%; Score 42; DB 7; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1117 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 19
US-10-425-405
; Sequence 405, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 405
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF095448
US-10-425-405

Query Match 100.0%; Score 42; DB 8; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1117 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 20
US-10-775-920-9
; Sequence 9, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-9

Query Match 100.0%; Score 42; DB 9; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||

Db 1117 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 21
US-10-510-507-2
; Sequence 2, Application US/10510507
; Publication No. US20050282165A1
; GENERAL INFORMATION:
; APPLICANT: Terrett, Jonathan A
; TITLE OF INVENTION: DIAGNOSIS OF CARCINOMA USING RAIG1 POLYPEPTIDES
; FILE REFERENCE: 2543-1-039PCT/US
; CURRENT APPLICATION NUMBER: US/10/510,507
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: GB0208331.9
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: GB0221538.2
; PRIOR FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-510-507-2

Query Match 100.0%; Score 42; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1117 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 22
US-10-775-920-12
; Sequence 12, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-12

Query Match 100.0%; Score 42; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||
Db 1121 GCCCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1162

RESULT 23
US-10-176-847-59
; Sequence 59, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST

```

; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-59

Query Match 100.0%; Score 42; DB 6; Length 2316;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1123 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1164

RESULT 24
US-11-080-991-59
; Sequence 59, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Pether Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: US/10/176,847
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-080-991-59

Query Match 100.0%; Score 42; DB 15; Length 2316;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1123 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1164

RESULT 25
US-10-775-920-11
; Sequence 11, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 00108
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR FILING DATE: US 60/447,900
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 11
; LENGTH: 2446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-11

```

```

Query Match 100.0%; Score 42; DB 9; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312

RESULT 26
US-10-225-567A-453
; Sequence 453, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 453
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-453

Query Match 100.0%; Score 42; DB 6; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312

RESULT 27
US-10-269-909-63
; Sequence 63, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303 (71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 63
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-63

Query Match 100.0%; Score 42; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 1312

```

Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 28

US-10-269-909-64
; Sequence 64, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUIZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-64

Query Match 100.0%; Score 42; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 29

US-10-295-027-619
; Sequence 619, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-01250005
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 619
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-619

Query Match 100.0%; Score 42; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 30

US-10-600-816-2
; Sequence 2, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-2

Query Match 100.0%; Score 42; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 31

US-10-600-816-18
; Sequence 18, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2

```
; SEQ ID NO 18
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RA13 Polymorphic Allele Summary Sequence.
;
; NAME/KEY: misc_feature
; LOCATION: (112)..(112)
; OTHER INFORMATION: wherein "n" equals either G or A.
;
; NAME/KEY: misc_feature
; LOCATION: (364)..(364)
; OTHER INFORMATION: wherein "n" equals either C or T.
;
; NAME/KEY: misc_feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: wherein "n" equals either C or T.
;
; NAME/KEY: misc_feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: wherein "n" equals either C or T.
;
; NAME/KEY: misc_feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: wherein "n" equals either A or G.
;
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: wherein "n" equals either A or G.
;
; NAME/KEY: misc_feature
; LOCATION: (1111)..(1111)
; OTHER INFORMATION: wherein "n" equals either T or C.
;
; NAME/KEY: misc_feature
; LOCATION: (1173)..(1173)
; OTHER INFORMATION: wherein "n" equals either A or G.
;
; US-10-600-816-18
;
; Query Match 100.0%; Score 42; DB 8; Length 2456;
; Best Local Similarity 100.0%; Pred. No. 1.1e-07;
; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
; 1271 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
;
; Db
;
; RESULT 32
; US-10-775-920-10
; Sequence 10, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; FILE REFERENCE: IN CERTAIN CANCERS
; CURRENT APPLICATION NUMBER: US/10/775,920
; PRIOR FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-10
;
; Query Match 100.0%; Score 42; DB 9; Length 2456;
; Best Local Similarity 100.0%; Pred. No. 1.1e-07;
; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; SEQ ID NO 18
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RA13 Polymorphic Allele Summary Sequence.
;
; NAME/KEY: misc_feature
; LOCATION: (112)..(112)
; OTHER INFORMATION: wherein "n" equals either G or A.
;
; NAME/KEY: misc_feature
; LOCATION: (364)..(364)
; OTHER INFORMATION: wherein "n" equals either C or T.
;
; NAME/KEY: misc_feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: wherein "n" equals either C or T.
;
; NAME/KEY: misc_feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: wherein "n" equals either C or T.
;
; NAME/KEY: misc_feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: wherein "n" equals either A or G.
;
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: wherein "n" equals either A or G.
;
; NAME/KEY: misc_feature
; LOCATION: (1111)..(1111)
; OTHER INFORMATION: wherein "n" equals either T or C.
;
; NAME/KEY: misc_feature
; LOCATION: (1173)..(1173)
; OTHER INFORMATION: wherein "n" equals either A or G.
;
; US-10-600-816-18
;
; Query Match 100.0%; Score 42; DB 8; Length 2456;
; Best Local Similarity 100.0%; Pred. No. 1.1e-07;
; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
; 1271 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
;
; Db
;
; RESULT 32
; US-10-775-920-10
; Sequence 10, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; FILE REFERENCE: IN CERTAIN CANCERS
; CURRENT APPLICATION NUMBER: US/10/775,920
; PRIOR FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-10
;
; Query Match 100.0%; Score 42; DB 9; Length 2456;
; Best Local Similarity 100.0%; Pred. No. 1.1e-07;
; Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


APPLICANT: Phillips, Heidi S.
APPLICANT: Polakis, Paul
APPLICANT: Spencer, Susan D.
APPLICANT: Smith, Victoria
APPLICANT: Williams, P. Mickey
APPLICANT: Wu, Thomas D.
APPLICANT: Zhang, Zemin
APPLICANT: Sakanaka, Chie
APPLICANT: Chuntharapai, Anan
APPLICANT: Reed Chae J.
TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
TITLE OF INVENTION: Treatment of Tumor
FILE REFERENCE: PS001R1PIB
CURRENT APPLICATION NUMBER: US/10/938,061
CURRENT FILING DATE: 2004-09-10
PRIOR FILING DATE: US 10/872,991
PRIOR APPLICATION NUMBER: US 10/872,991
PRIOR FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: US 10/872,972
PRIOR FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: US 10/241,220
PRIOR FILING DATE: 2002-09-11
PRIOR APPLICATION NUMBER: US 10/177,488
PRIOR FILING DATE: 2002-06-19
PRIOR APPLICATION NUMBER: US 60/299,500
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 60/301,880
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/323,268
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US 60/557,116
PRIOR FILING DATE: 2004-03-26
PRIOR APPLICATION NUMBER: US 60/598,899
PRIOR FILING DATE: 2004-08-04
NUMBER OF SEQ ID NOS: 154
SEQ ID NO 40
LENGTH: 2456
TYPE: DNA
ORGANISM: Homo sapiens
US-10-938-061-40

Query Match 100.0%; Score 42; DB 10; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 35
US-11-169-041-32
Sequence 32, Application US/11169041
Publication No. US20060019284A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
TITLE OF INVENTION: CELLS
FILE REFERENCE: 10001 NP
CURRENT APPLICATION NUMBER: US/11/169,041
CURRENT FILING DATE: 2005-06-28
PRIOR APPLICATION NUMBER: 60/584,405
PRIOR FILING DATE: 2004-06-30
NUMBER OF SEQ ID NOS: 527
SOFTWARE: PatentIn version 3.2
SEQ ID NO 32
LENGTH: 2456
TYPE: DNA
ORGANISM: Homo sapiens
US-11-169-041-32

Query Match 100.0%; Score 42; DB 16; Length 2456;

Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1271 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 36
US-10-264-049-834
Sequence 834, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 834
LENGTH: 2593
TYPE: DNA
ORGANISM: Homo sapiens
US-10-264-049-834

Query Match 100.0%; Score 42; DB 7; Length 2593;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1383 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1424

RESULT 37
US-10-198-846-10424
Sequence 10424, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10424
LENGTH: 4239
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1, 4238, 4239
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10424

Query Match 100.0%; Score 42; DB 6; Length 4239;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42

Db 1679 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1720
|||||

RESULT 38

US-10-505-486-196
; Sequence 196, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0008PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 196
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Human
US-10-505-486-196

Query Match 96.2%; Score 40.4; DB 10; Length 1788;
Best Local Similarity 97.6%; Pred. No. 5.3e-07;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||

Db 1018 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1059

RESULT 39

US-10-600-816-33
; Sequence 33, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RA13, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate Oligonucleotide.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)..(39)
; OTHER INFORMATION: wherein "n" equals A, C, G, or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6)..(33)
; OTHER INFORMATION: wherein "y" equals C, or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (27)..(42)
; OTHER INFORMATION: wherein "r" equals A, or G.
US-10-600-816-33

Query Match 71.4%; Score 30; DB 8; Length 42;
Best Local Similarity 64.3%; Pred. No. 0.0081;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
|||||

Db 1 GCNCAYGNTGGCCNTCCNTCNCNTAYAAAGATTAYGARGTNAAR 42
|||||

RESULT 40

US-11-060-756-43361
; Sequence 43361, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM01083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43361
; LENGTH: 25
; TYPE: DNA
; ORGANISM: probe
US-11-060-756-43361

Query Match 59.5%; Score 25; DB 13; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTA 39
|||||

Db 1 GAGCCCTTACAAAGACTATGAAGTA 25
|||||

RESULT 41

US-10-459-155-7
; Sequence 7, Application US/10459155
; Publication No. US20040254365A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Sho Tone
; TITLE OF INVENTION: IMMUNIZATION AGAINST FLAVIVIRUS
; FILE REFERENCE: 08919-101001
; CURRENT APPLICATION NUMBER: US/10/459,155
; CURRENT FILING DATE: 2003-06-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Japanese encephalitis virus
US-10-459-155-7

Query Match 54.3%; Score 22.8; DB 9; Length 480;
Best Local Similarity 79.4%; Pred. No. 21;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CAGCGTTGGCGAGCCCTTACAAAGACTATGAAG 37
|||||

Db 330 CAGCGTTGGCGAGCCCTTTCACACGACTTTGAAG 363
|||||

RESULT 42

US-10-459-155-12
; Sequence 12, Application US/10459155
; Publication No. US20040254365A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Sho Tone
; TITLE OF INVENTION: IMMUNIZATION AGAINST FLAVIVIRUS

```
; FILE REFERENCE: 08919-101001
; CURRENT APPLICATION NUMBER: US/10/459,155
; CURRENT FILING DATE: 2003-06-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated oligonucleotide
US-10-459-155-12

Query Match          54.3%; Score 22.8; DB 9; Length 525;
Best Local Similarity 79.4%; Pred. No. 21;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
      ||||| ||| ||||| ||| ||||| |||||
Db  375 CACGCTGGCAAGCCCTTTTCAACGACTTTGAAG 408

RESULT 43
US-10-459-155-9
; Sequence 9, Application US/10459155
; Publication No. US20040254365A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Sho Tone
; TITLE OF INVENTION: IMMUNIZATION AGAINST FLAVIVIRUS
; FILE REFERENCE: 08919-101001
; CURRENT APPLICATION NUMBER: US/10/459,155
; CURRENT FILING DATE: 2003-06-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Japanese encephalitis virus
US-10-459-155-9

Query Match          54.3%; Score 22.8; DB 9; Length 1500;
Best Local Similarity 79.4%; Pred. No. 28;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
      ||||| ||| ||||| ||| ||||| |||||
Db  1203 CACGCTGGCAAGCCCTTTTCAACGACTTTGAAG 1236

RESULT 44
US-10-459-155-21
; Sequence 21, Application US/10459155
; Publication No. US20040254365A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Sho Tone
; TITLE OF INVENTION: IMMUNIZATION AGAINST FLAVIVIRUS
; FILE REFERENCE: 08919-101001
; CURRENT APPLICATION NUMBER: US/10/459,155
; CURRENT FILING DATE: 2003-06-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Japanese encephalitis virus
US-10-459-155-21

Query Match          54.3%; Score 22.8; DB 9; Length 1545;
Best Local Similarity 79.4%; Pred. No. 28;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
      ||||| ||| ||||| ||| ||||| |||||
Db  1203 CACGCTGGCAAGCCCTTTTCAACGACTTTGAAG 1236

; FILE REFERENCE: 08919-101001
; CURRENT APPLICATION NUMBER: US/10/459,155
; CURRENT FILING DATE: 2003-06-11
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated oligonucleotide
US-10-459-155-12

Query Match          54.3%; Score 22.8; DB 9; Length 525;
Best Local Similarity 79.4%; Pred. No. 21;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  4 CACGCTTGGCGAGCCCTTACAAAGACTATGAAG 37
      ||||| ||| ||||| ||| ||||| |||||
Db  375 CACGCTGGCAAGCCCTTTTCAACGACTTTGAAG 408

RESULT 45
US-09-925-065A-691738
; Sequence 691738, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 691738
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-691738

Query Match          53.3%; Score 22.4; DB 4; Length 1540;
Best Local Similarity 72.5%; Pred. No. 41;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy  3 CCAGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
      ||||| ||| ||||| ||| ||||| |||||
Db  870 CAAACTAGACCAAGGCTTTTACAAAAAAGTAAAGTATAG 909

RESULT 46
US-09-925-065A-691739
; Sequence 691739, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 691739
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-691739

Query Match          53.3%; Score 22.4; DB 4; Length 1540;
Best Local Similarity 72.5%; Pred. No. 41;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```



```

; ORGANISM: Homo sapiens
US-09-925-065A-691740

Query Match 53.3%; Score 22.4; DB 5; Length 1540;
Best Local Similarity 72.5%; Pred. No. 41;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 3 CCAGCTTGCCGCGACCTTACAAAGACTATGAAGTAAAG 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 870 CAAAAGTAGACCAAGGCTTTACAAAAAATGAAGATATAG 909

RESULT 51
US-09-917-800A-76
; Sequence 76, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA618700
US-09-917-800A-76

Query Match 52.9%; Score 22.2; DB 3; Length 584;
Best Local Similarity 88.9%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 15 GAGCCCTTACAAAGACTATGAAGTAA 41
   ||| ||| ||| ||| ||| ||| ||| |||
Db 373 GAGCCCTTACAAAGACTATGAAGCAAA 399

RESULT 52
US-09-925-065A-49510
; Sequence 49510, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135

```

```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150748
; LENGTH: 2548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-150748

Query Match          52.9%; Score 22.2; DB 12; Length 2548;
Best Local Similarity 77.1%; Pred. No. 57;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
   |||||
Db 374 GCCTTGACAAGTTATTACAAAGAATATGAAGTAAA 408

RESULT 55
US-10-301-480-764157
; Sequence 764157, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 764157
; LENGTH: 2548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-764157

Query Match          52.9%; Score 22.2; DB 12; Length 2548;
Best Local Similarity 77.1%; Pred. No. 57;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
   |||||
Db 374 GCCTTGACAAGTTATTACAAAGAATATGAAGTAAA 408

RESULT 56
US-10-017-161-455/c
; Sequence 455, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ASURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; - CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18

```



```
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 126
; LENGTH: 4826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 238854.23
US-10-044-090-126

Query Match      51.9%; Score 21.8; DB 6; Length 4826;
Best Local Similarity 70.7%; Pred. No. 1e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 41
Db 4091 GCGAAGTTTCACAGAGTCTTCAAAAGACTCTGAAGAAA 4051

RESULT 59
US-10-084-817-304/c
; Sequence 304, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 2002-02-25
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 304
; LENGTH: 4826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 238854.23
US-10-084-817-304

Query Match      51.9%; Score 21.8; DB 6; Length 4826;
Best Local Similarity 70.7%; Pred. No. 1e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 41
Db 4091 GCGAAGTTTCACAGAGTCTTCAAAAGACTCTGAAGAAA 4051

RESULT 60
US-09-764-847-1315
; Sequence 1315, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1315
; LENGTH: 17849
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-764-847-1315

Query Match      51.9%; Score 21.8; DB 3; Length 17849;
Best Local Similarity 70.7%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 13871 CCCCACCTTCTAAGAGGCTCTTATAAAAACTTTCAAGTAA 13911

RESULT 61
US-09-764-891-10139
; Sequence 10139, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10139
; LENGTH: 17849
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-764-891-10139

Query Match      51.9%; Score 21.8; DB 3; Length 17849;
Best Local Similarity 70.7%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 13871 CCCCACCTTCTAAGAGGCTCTTATAAAAACTTTCAAGTAA 13911

RESULT 62
US-10-092-154-1315
; Sequence 1315, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1315
; LENGTH: 17849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1315

Query Match      51.9%; Score 21.8; DB 6; Length 17849;
Best Local Similarity 70.7%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAA 42
Db 13871 CCCCACCTTCTAAGAGGCTCTTATAAAAACTTTCAAGTAA 13911

RESULT 63
US-09-764-847-1313
; Sequence 1313, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```


Query Match	51.4%	Score 21.6;	DB 9;	Length 70000;
Best Local Similarity	75.0%	Pred. NO. 2.3e+02;		
Matches 27; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;

Qy 1 G C C C A C G C T T G G C G A G C C T T A C A A G A G A C T A T G A A 36
 | | | | | | | | | | | | | | | | | | | |
Db 42352 G T C C A G G C T T G G T C A G C C T T A G A A G G G C T A G C A A 42317

```

RESULT 68
US-11-097-143-1964
; Sequence 1964, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1964
; LENGTH: 2356
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-1964

```

```
Query Match      51.0%; Score 21.4; DB 13; Length 2356;
Best Local Similarity 71.8%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

Qy 1 GCCACGCTTGGCCGAGCCTTACAAAGACTATGAAGTA 39
||||| | | | | | | | | | | | | | |
Db 1212 GCCAAGAATCGCAGAGCCATTATCAAGAAATATCAAGGA 1250

```

RESULT 69
US-11-096-568A-26327
; Sequence 26327, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26327
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (1)..(1227)
; OTHER INFORMATION: Ceres Seq. ID no. 13501189
US-11-096-568A-26327

```

```

Query Match      50.5%; Score 21.2; DB 16; Length 1227;
Best Local Similarity 69.0%; Pred. No. 1.3e+02;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0

```

```

RESULT 70
US-10-450-763-28218/c
; Sequence 28218, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 28218
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (651)..(1361)
; OTHER INFORMATION: 100% homologous to Homo sapiens dJ185D5.1 (novel
; OTHER INFORMATION: gene), accession number AL118498, Smith-Waterman Score=1257.
US-10-450-763-28218

```

Query Match	50.5%	Score 21.2;	DB 10;	Length 1446;
Best Local Similarity	76.5%;	Pred. No. 1.4e+02;		
Matches 26;	Conservative	0;	Mismatches 8;	Indels 0;
Gaps	0;			

Qy 1 G C C A C G C T T G G C G A G C C C T T A C A A G A C T A T G 34
 ||| ||| ||| ||| ||| ||| ||| |||
Db 204 G C C T T C G C T C G G C G A G C C T C A G G A A G A C T A T G 171

```

RESULT 71
US-10-369-493-36568/c
; Sequence 36568, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICRO
; TITLE OF INVENTION: PLANTS WITH IMPRO
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/3650, 03
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36568
; LENGTH: 1555
; TYPE: DNA

```

```
; ORGANISM: Aspergillus nidulans
US-10-369-493-36568

Query Match          50.5%; Score 21.2; DB 7; Length 1555;
Best Local Similarity 76.5%; Pred. No. 1.4e+02;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCACGCTTGGCGAGCCCTTACAAAGACTATGA 35
   ||| ||||| ||||| ||||| ||||| |||||
Db 299 CCAACATTGGCGTACCTTACGCTGACTATGA 266

RESULT 72
US-10-425-115-130787/c
; Sequence 130787, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihwei
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 130787
; LENGTH: 3141
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_50754C.1
US-10-425-115-130787

Query Match          50.5%; Score 21.2; DB 9; Length 3141;
Best Local Similarity 69.0%; Pred. No. 1.7e+02;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 42
   ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1297 GCTTAGCATGGCTGAGCCATACACAGATGTTGATGACAG 1256

RESULT 73
US-09-814-353-3621
; Sequence 3621, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3621
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 512_519
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-9935

Query Match          50.0%; Score 21; DB 3; Length 527;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
   ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 458 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAAA 494

RESULT 75
US-09-814-353-5769
; Sequence 5769, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
```

```
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 512_519
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-3621

Query Match          50.0%; Score 21; DB 3; Length 527;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
   ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 458 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAAA 494

RESULT 74
US-09-814-353-9935
; Sequence 9935, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9935
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 512_519
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-9935

Query Match          50.0%; Score 21; DB 3; Length 527;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 ACCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
   ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 458 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAAA 494

RESULT 75
US-09-814-353-5769
; Sequence 5769, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
```

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5769
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 514
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-5769

Query Match 50.0%; Score 21; DB 3; Length 559;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | | |
Db 458 AAGTTTCACAGAGTCCTTCAAAGAGACTCTGAAGAAA 494

RESULT 76
US-09-814-353-12052
; Sequence 12052, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12052
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: 514
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-12052

Query Match 50.0%; Score 21; DB 3; Length 559;
Best Local Similarity 73.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | | |
Db 458 AAGTTTCACAGAGTCCTTCAAAGAGACTCTGAAGAAA 494

RESULT 77
US-09-814-353-4193
; Sequence 4193, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4193
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 515
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4193

Query Match 50.0%; Score 21; DB 3; Length 579;
Best Local Similarity 73.0%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | | |
Db 459 AAGTTTCACAGAGTCCTTCAAAGAGACTTTGAAGAAA 495

RESULT 78
US-09-814-353-10499
; Sequence 10499, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 22037
; SEQ ID NO 10499
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 515
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-10499
```

```
Query Match 50.0%; Score 21; DB 3; Length 579;
Best Local Similarity 73.0%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | |
Db 459 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAA 495
```

RESULT 79

```
US-09-814-353-18436
; Sequence 18436, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18436
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 606
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-18436
```

```
Query Match 50.0%; Score 21; DB 3; Length 625;
```

```
Best Local Similarity 73.0%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | |
Db 538 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAA 574
```

RESULT 80

```
US-09-814-353-18787/c
; Sequence 18787, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18787
; LENGTH: 692
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-18787
```

```
Query Match 50.0%; Score 21; DB 3; Length 692;
Best Local Similarity 73.0%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
Qy 5 ACGCTTGGCGAGCCCTTACAAAGACTATGAAGTAAA 41
| | | | | | | | | | | | | | | | | | | | |
Db 521 AAGTTTCACAGAGTCCTTCAAAAGACTTTGAAGAAA 485
```

RESULT 81

```
US-09-814-353-20633
; Sequence 20633, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
```


; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20633
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 1096
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20633

Query Match 50.0%; Score 21; DB 3; Length 1096;
Best Local Similarity 73.0%; Pred. No. 1.6e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
Db 437 AAGTTTCACAGAGCTCTTCAAAGAGACTCTGAAGAAA 473

RESULT 82

; Sequence 20810, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20810
; LENGTH: 1386
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1383, 1384, 1385, 1386
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20810

Query Match 50.0%; Score 21; DB 3; Length 1386;
Best Local Similarity 73.0%; Pred. No. 1.7e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
Db 566 AAGTTTCACAGAGCTCTTCAAAGAGACTCTGAAGAAA 602

RESULT 83

US-09-814-353-20790
; Sequence 20790, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20790
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 2114, 2130, 2138, 2144, 2145
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20790

Query Match 50.0%; Score 21; DB 3; Length 2145;
Best Local Similarity 73.0%; Pred. No. 1.8e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41
Db 657 AAGTTTCACAGAGCTCTTCAAAGAGACTCTGAAGAAA 693

RESULT 84

US-10-006-285-420/c
; Sequence 420, Application US/10006285
; Publication No. US20030165854A1
; GENERAL INFORMATION:
; APPLICANT: Mary Jane Cunningham
; APPLICANT: Matthew R. Kaser
; TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS
; FILE REFERENCE: PA-0039 US

; CURRENT APPLICATION NUMBER: US/10/006,285
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 514
; SOFTWARE: PERL Program
; SEQ ID NO 420
; LENGTH: 4605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165854A1 238854.22
US-10-006-285-420

Query Match 50.0%; Score 21; DB 7; Length 4605;
Best Local Similarity 73.0%; Pred. No. 2.2e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 ACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAA 41

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
||||| ||||| ||||| ||||| ||||| |||||
Db 35 CCCAACCTTGGCAGAGCCCATCACAAGTACTAT 4

RESULT 89

US-09-925-065A-148760/c
; Sequence 148760, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148760
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-148760

Query Match 49.5%; Score 20.8; DB 5; Length 560;
Best Local Similarity 78.1%; Pred. No. 1.6e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
||||| ||||| ||||| ||||| ||||| |||||
Db 35 CCCAACCTTGGCAGAGCCCATCACAAGTACTAT 4

RESULT 90

US-09-925-065A-148761/c
; Sequence 148761, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148761
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-148761

Query Match 49.5%; Score 20.8; DB 5; Length 560;
Best Local Similarity 78.1%; Pred. No. 1.6e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
||||| ||||| ||||| ||||| ||||| |||||
Db 35 CCCAACCTTGGCAGAGCCCATCACAAGTACTAT 4

RESULT 91

US-10-301-480-242891/c
; Sequence 242891, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/301.480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242891
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-242891

Query Match 49.5%; Score 20.8; DB 12; Length 569;
Best Local Similarity 78.1%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
||||| ||||| ||||| ||||| ||||| |||||
Db 44 CCCAACCTTGGCAGAGCCCATCACAAGTACTAT 13

RESULT 92

US-10-301-480-242892/c
; Sequence 242892, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/301.480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242892
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-242892

Query Match 49.5%; Score 20.8; DB 12; Length 569;
Best Local Similarity 78.1%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
||||| ||||| ||||| ||||| ||||| |||||
Db 44 CCCAACCTTGGCAGAGCCCATCACAAGTACTAT 13

```
RESULT 93
US-10-301-480-856300/c
; Sequence 856300, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 856300
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-856300
Query Match 49.5%; Score 20.8; DB 12; Length 569;
Best Local Similarity 78.1%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
Db 44 CCCACCTTGGCAGAGCCATCACAAAGTACTAT 13
RESULT 94
US-10-301-480-856301/c
; Sequence 856301, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 856301
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-856301
Query Match 49.5%; Score 20.8; DB 12; Length 569;
Best Local Similarity 78.1%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CCCACGCTTGGCCGAGCCCTTACAAAGACTAT 33
Db 44 CCCACCTTGGCAGAGCCATCACAAAGTACTAT 13
RESULT 95
US-09-925-065A-716598/c
; Sequence 716598, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 716598
; LENGTH: 1465
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-716598
Query Match 49.5%; Score 20.8; DB 4; Length 1465;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 3 CCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1184 CCGTGCTTGGCCGAGATCCTATAATGACTAGTAGTATGGAGAG 1145
RESULT 96
US-09-925-065A-716599/c
; Sequence 716599, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 716599
; LENGTH: 1465
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-716599
Query Match 49.5%; Score 20.8; DB 4; Length 1465;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 3 CCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 42
Db 1184 CCGTGCTTGGCCGAGATCCTATAATGACTAGTAGTATGGAGAG 1145
RESULT 97
US-09-925-065A-716600/c
; Sequence 716600, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

US-09-925-065A-716599/c
; Sequence 716599, Application US/09925065A

Search completed: June 10, 2006, 20:50:45
Job time : 1636.6 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:18:44 ; Search time 463.533 Seconds
(without alignments)

2345.260 Million cell updates/sec

Title: US-10-600-816-31

Perfect score: 17

Sequence: 1 gccacgctggccgag 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sv.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	17	100.0	603	2	BD233463 Human pro
2	17	100.0	1212	2	BD209699 Compositi
3	17	100.0	1212	2	AX341505 Sequence
4	17	100.0	1619	2	AX078375 Sequence
5	17	100.0	1718	2	BD233473 Human pro
6	17	100.0	1826	5	AK172760 Homo sapi
7	17	100.0	2290	2	C0723177 Sequence
8	17	100.0	2296	5	BC003665 Homo sapi
9	17	100.0	2297	5	AF506289 Homo sapi
10	17	100.0	2302	2	AX930411 Sequence
11	17	100.0	2302	5	AF095448 Homo sapi
12	17	100.0	2446	2	BD156680 Primer fo
13	17	100.0	2446	2	AX877483 Sequence
14	17	100.0	2446	5	AK001761 Homo sapi
15	17	100.0	2456	2	CQ981495 Sequence
16	17	100.0	2456	2	BD210040 Methods o
17	17	100.0	2456	2	AX549168 Sequence
18	17	100.0	3057	5	AK122672 Homo sapi

19	17	100.0	3371	2	AX188348 Sequence
20	17	100.0	6730	2	CQ894732 Sequence
c 21	17	100.0	161577	5	AC007688 Homo sapi
c 22	17	100.0	213601	12	AC165037 Bos tauru
23	17	100.0	284106	12	AC166087 Bos tauru
c 24	17	100.0	299050	15	SC039104 Streptomy
25	16	94.1	110000	15	CP000124_29 Continuation (30 o
26	16	94.1	110000	15	BX571965_26 Continuation (27 o
27	16	94.1	110000	15	CP000010_16 Continuation (17 o
c 28	16	94.1	174298	12	AC167299 Oryctolag
c 29	15.4	90.6	117	15	MAU564400 Mycobacte
30	15.4	90.6	385	5	AM181332 Homo sapi
c 31	15.4	90.6	586	1	AY535046 Unculture
c 32	15.4	90.6	738	10	DQ288979 Bovine le
33	15.4	90.6	1802	1	AY381191 Unculture
c 34	15.4	90.6	2306	2	AX416413 Sequence
35	15.4	90.6	2550	5	HSA564991 Homo sapi
36	15.4	90.6	2550	5	HSA564992 Homo sapi
37	15.4	90.6	2550	5	HSA564993 Homo sapi
38	15.4	90.6	2623	5	BC018939 Homo sapi
39	15.4	90.6	2623	5	BC034926 Homo sapi
40	15.4	90.6	3252	4	AF420019 Aspergill
41	15.4	90.6	3325	2	CQ727552 Sequence
42	15.4	90.6	3602	5	AF542080 Homo sapi
43	15.4	90.6	4183	2	AR157945 Sequence
44	15.4	90.6	4183	2	BD103585 Human nuc
45	15.4	90.6	4232	2	CQ850573 Sequence
46	15.4	90.6	4232	5	AK127735 Homo sapi
47	15.4	90.6	4641	5	AB205580 Homo sapi
48	15.4	90.6	5425	5	AY375870 Homo sapi
49	15.4	90.6	5599	5	AY375871 Homo sapi
50	15.4	90.6	5600	5	AY375872 Homo sapi
51	15.4	90.6	5600	5	AY375873 Homo sapi
c 52	15.4	90.6	8714	10	BLVCG
c 53	15.4	90.6	10694	15	AE004474 Pseudomon
54	15.4	90.6	16651	4	AB168089 Chlamydom
55	15.4	90.6	59316	6	BX000454 Mouse DNA
56	15.4	90.6	66661	12	AC165982 Bos tauru
57	15.4	90.6	69966	5	AY663411 Homo sapi
58	15.4	90.6	100761	5	AY663415 Homo sapi
c 59	15.4	90.6	103216	5	HS039318 Human DNA
60	15.4	90.6	103395	5	AY663406 Homo sapi
61	15.4	90.6	104996	5	AY663395 Homo sapi
c 62	15.4	90.6	110000	15	AE001439_02 Continuation (3 of
c 63	15.4	90.6	110000	15	AE001439_03 Continuation (4 of
c 64	15.4	90.6	110000	15	CP000240_27 Continuation (28 o
c 65	15.4	90.6	110000	15	CP000240_28 Continuation (29 o
c 66	15.4	90.6	110000	15	CP000248_28 Continuation (9 of
c 67	15.4	90.6	110000	15	AE014295_08 Continuation (14 o
c 68	15.4	90.6	110000	15	AE017262_13 Continuation (13 o
c 69	15.4	90.6	110000	15	BA000032_12 AP003159 Homo sapi
c 70	15.4	90.6	113733	5	AL662789 Human DNA
c 71	15.4	90.6	147557	5	AL662789 Human DNA
c 72	15.4	90.6	150372	12	AC150103 Gallus ga
73	15.4	90.6	150375	12	AC169542 Bos tauru
74	15.4	90.6	153108	5	AL355273 Human DNA
c 75	15.4	90.6	160714	12	AC150149 Gallus ga
76	15.4	90.6	185122	5	AC011448 Homo sapi
77	15.4	90.6	188122	12	AC146997 Arbacia p
c 78	15.4	90.6	172995	6	AC130666 Mus muscu
c 79	15.4	90.6	179959	5	AL160395 Human DNA
c 80	15.4	90.6	183887	5	AC017033 Homo sapi
81	15.4	90.6	187027	12	AC019078 Homo sapi
c 82	15.4	90.6	193015	12	AC150074 Gallus ga
83	15.4	90.6	195269	2	AX417035 Sequence
c 84	15.4	90.6	196542	5	AC002365 Homo sapi
c 85	15.4	90.6	197456	12	AC005054 Homo sapi
c 86	15.4	90.6	198375	6	AC120873 Mus muscu
c 87	15.4	90.6	200618	5	AC009061 Homo sapi
88	15.4	90.6	205915	6	AC119828 Mus muscu
c 89	15.4	90.6	209262	12	AC021091 Homo sapi
90	15.4	90.6	214798	6	AC091522 Mus muscu
91	15.4	90.6	216540	12	AC134138 Rattus no

92	15.4	90.6	218303	6	CR974487	CR974487 Mouse DNA	c 165	14.4	84.7	14402	15	AF098795	AF098795 Pseudomon
c 93	15.4	90.6	230993	12	AC150072	AC150072 Gallus ga	c 166	14.4	84.7	37294	5	HSINPPI2	AF141325 Homo sapi
94	15.4	90.6	233345	12	AC150066	AC150066 Gallus ga	c 167	14.4	84.7	41921	2	AY341028	AY341028 Homo sapi
95	15.4	90.6	234079	12	AC130741	AC130741 Rattus no	c 168	14.4	84.7	48169	2	CS184722	CS184722 Sequence
96	15.4	90.6	236105	12	AC155587	AC155587 Bos tauru	c 169	14.4	84.7	48169	15	AJ628955	AJ628955 Streptomy
c 97	15.4	90.6	240430	12	AC136419	AC136419 Rattus no	c 170	14.4	84.7	49312	2	AR284435	AR284435 Sequence
98	15.4	90.6	241836	12	AC106342	AC106342 Rattus no	c 171	14.4	84.7	64650	15	AM113472	AM113472 Micromono
c 99	15.4	90.6	245117	12	AC111284	AC111284 Rattus no	c 172	14.4	84.7	70692	4	HSAP07155_41	HSAP07155_41 o
c 100	15.4	90.6	247212	12	AC115439	AC115439 Rattus no	c 173	14.4	84.7	74517	5	HSAP09615	HSAP09615 Homo sapi
c 101	15.4	90.6	250050	15	AL591978	AL591978 Listeria	c 174	14.4	84.7	77878	12	AC126362	AC126362 Homo sapi
c 102	15.4	90.6	256747	12	AC125838	AC125838 Rattus no	c 175	14.4	84.7	81837	5	AC004775	AC004775 Homo sapi
c 103	15.4	90.6	266814	12	AC129165	AC129165 Rattus no	c 176	14.4	84.7	82412	12	AC040940	AC040940 Homo sapi
c 104	15.4	90.6	333050	15	AL596168	AL596168 Listeria	c 177	14.4	84.7	83888	12	AP006735	AP006735 Lotus cor
c 105	15.4	90.6	349980	2	AX417044	AX417044 Sequence	c 178	14.4	84.7	88737	15	AP006878	AP006878 (21 o
c 106	15.4	90.6	349980	2	AX492784	AX492784 Sequence	c 179	14.4	84.7	91649	4	AP006737	AP006737 Lotus cor
c 107	15.4	90.6	349980	2	AX492785	AX492785 Sequence	c 180	14.4	84.7	92375	4	AP003905	AP003905 Oryza sat
c 108	15.4	90.6	349980	2	AX553951	AX553951 Sequence	c 181	14.4	84.7	92564	5	AY007685	AY007685 Homo sapi
c 109	15.4	90.6	349980	2	AX553952	AX553952 Sequence	c 182	14.4	84.7	99967	4	NCB33B10	NCB33B10 Neospor
c 110	15.4	90.6	349980	2	AX641668	AX641668 Sequence	c 183	14.4	84.7	105040	5	AC011229	AC011229 Homo sapi
c 111	15	88.2	337	2	CQ525744	CQ525744 Sequence	c 184	14.4	84.7	110000	4	AP008218	AP008218 156
c 112	15	88.2	10029	15	AE009464	AE009464 Brucella	c 185	14.4	84.7	110000	4	AP008218	AP008218 170
c 113	15	88.2	11236	15	AE012464	AE012464 Xanthomon	c 186	14.4	84.7	110000	4	AP008218	AP008218 171
c 114	15	88.2	110000	15	CP000050_08	Continuation (9 of	c 187	14.4	84.7	110000	4	CR382128	CR382128 26
c 115	15	88.2	110000	15	CP000058_14	Continuation (15 o	c 188	14.4	84.7	110000	4	AP007171	AP007171 17
c 116	15	88.2	110000	15	CP000115_23	Continuation (24 o	c 189	14.4	84.7	110000	4	AP008207	AP008207 102
c 117	15	88.2	110000	15	AE014291_17	Continuation (18 o	c 190	14.4	84.7	110000	4	AP008210	AP008210 338
c 118	15	88.2	110000	15	AE017223_17	Continuation (18 o	c 191	14.4	84.7	110000	4	AP008210	AP008210 339
c 119	15	88.2	110000	15	AM040264_17	Continuation (18 o	c 192	14.4	84.7	110000	12	CR954200	CR954200 00
c 120	15	88.2	161423	12	AC151882	AC151882 Sus scrof	c 193	14.4	84.7	110000	12	CR954202	CR954202 Oestreococ
c 121	15	88.2	178648	12	AC156675	AC156675 Bos tauru	c 194	14.4	84.7	110000	12	CR954202	CR954202 Oestreococ
c 122	14.4	84.7	344	2	AX245499	AX245499 Sequence	c 195	14.4	84.7	110000	15	CP000075	CP000075 36
c 123	14.4	84.7	349	5	HUMHDDQ802	M27657 Homo sapien	c 196	14.4	84.7	110000	15	CP000090	CP000090 Ralstonia
c 124	14.4	84.7	513	2	CQ742410	CQ742410 Sequence	c 197	14.4	84.7	110000	15	CP000092	CP000092 3
c 125	14.4	84.7	539	4	AK105321	AK105321 Oryza sat	c 198	14.4	84.7	110000	15	CP000110	CP000110 Continuation (4 of
c 126	14.4	84.7	546	2	BD151780	BD151780 Primer fo	c 199	14.4	84.7	110000	15	CP000124	CP000124 Continuation (10 of
c 127	14.4	84.7	546	2	AX871718	AX871718 Sequence	c 200	14.4	84.7	110000	15	CP000230	CP000230 Continuation (7 of
c 128	14.4	84.7	550	1	AM160719	AM160719 Unidentif	c 201	14.4	84.7	110000	15	CP000248	CP000248 25
c 129	14.4	84.7	570	2	CQ522174	CQ522174 Sequence	c 202	14.4	84.7	110000	15	CP000249	CP000249 21
c 130	14.4	84.7	572	7	BV244429	BV244429 S234P6330	c 203	14.4	84.7	110000	15	CP000250	CP000250 30
c 131	14.4	84.7	650	7	BV589117	BV589117 G591P6371	c 204	14.4	84.7	110000	15	CR555306	CR555306 03
c 132	14.4	84.7	680	2	BD150346	BD150346 Primer fo	c 205	14.4	84.7	110000	15	AE016822	AE016822 04
c 133	14.4	84.7	680	2	AX870284	AX870284 Sequence	c 206	14.4	84.7	110000	15	AE016853	AE016853 52
c 134	14.4	84.7	686	7	BV561948	BV561948 qie74H05.	c 207	14.4	84.7	110000	15	AY596297	AY596297 28
c 135	14.4	84.7	703	7	BV620547	BV620547 S217P6235	c 208	14.4	84.7	110000	15	BA000022	BA000022 01
c 136	14.4	84.7	706	7	BV586726	BV586726 G591P6333	c 209	14.4	84.7	110000	15	BA000039	BA000039 21
c 137	14.4	84.7	765	7	BV610738	BV610738 S216P6152	c 210	14.4	84.7	110000	15	BA000039	BA000039 22
c 138	14.4	84.7	768	15	AB182132	AB182132 Pseudomon	c 211	14.4	84.7	110000	15	BA000039	BA000039 22
c 139	14.4	84.7	781	7	BV532865	BV532865 G591P6196	c 212	14.4	84.7	110000	15	EX571965	EX571965 07
c 140	14.4	84.7	789	7	BV550427	BV550427 S215P6001	c 213	14.4	84.7	110000	15	EX363398	EX363398 39
c 141	14.4	84.7	825	7	BV591270	BV591270 G591P6411	c 214	14.4	84.7	110000	15	CP000010	CP000010 02
c 142	14.4	84.7	899	7	BV570729	BV570729 G591P6419	c 215	14.4	84.7	115339	4	AP003940	AP003940 Oryza sat
c 143	14.4	84.7	920	7	BV537670	BV537670 G591P6382	c 216	14.4	84.7	122576	4	NCB13H18	NCB13H18 Neurospor
c 144	14.4	84.7	934	7	BV578148	BV578148 G591P6107	c 217	14.4	84.7	126910	5	AC026794	AC026794 Homo sapi
c 145	14.4	84.7	992	4	VD302674	VD302674 Verticill	c 218	14.4	84.7	127477	5	BS000128	BS000128 Pan trogl
c 146	14.4	84.7	1194	1	AY046706	AY046706 Unculture	c 219	14.4	84.7	128073	13	AY321476	AY321476 Tribolium
c 147	14.4	84.7	1197	1	AY046676	AY046676 Unculture	c 220	14.4	84.7	132488	5	AC007963	AC007963 Homo sapi
c 148	14.4	84.7	1197	1	AY046690	AY046690 Sequence	c 221	14.4	84.7	132488	4	CNS07Y99	CNS07Y99
c 149	14.4	84.7	1231	2	CQ723659	CQ723659 Sequence	c 222	14.4	84.7	136254	4	AC011036	AC011036 Homo sapi
c 150	14.4	84.7	1269	15	SCF8P33	SCF8P33 S.chrysonal	c 223	14.4	84.7	138129	12	AC011036	AC011036 Homo sapi
c 151	14.4	84.7	1399	5	BC013935	BC013935 Homo sapi	c 224	14.4	84.7	140572	5	AL354743	AL354743 Human DNA
c 152	14.4	84.7	1536	15	HPAMIE	Y12552 H.pylori am	c 225	14.4	84.7	141056	5	AL354836	AL354836 Macaca mu
c 153	14.4	84.7	1788	4	AC080310	AC080310 Trebouxia	c 226	14.4	84.7	141130	12	AC142717	AC142717 Dasygyn n
c 154	14.4	84.7	1789	4	TM18SRRNA	Z21552 T.magna gen	c 227	14.4	84.7	147774	12	AC156758	AC156758 Dasygyn n
c 155	14.4	84.7	2489	5	BC050058	BC050058 Homo sapi	c 228	14.4	84.7	148328	6	AL808120	AL808120 Pongo pyg
c 156	14.4	84.7	2537	5	AY816326	AY816326 Homo sapi	c 229	14.4	84.7	148504	5	AC138732	AC138732 Pongo pyg
c 157	14.4	84.7	2662	14	AF071799	AF071799 Delphinap	c 230	14.4	84.7	149463	6	AC137096	AC137096 Mus muscu
c 158	14.4	84.7	3020	6	AF046907	AF046907 Mus muscu	c 231	14.4	84.7	152680	12	AC161715	AC161715 Sorex ara
c 159	14.4	84.7	3657	5	BC021123	BC021123 Homo sapi	c 232	14.4	84.7	157633	12	AC137826	AC137826 Medicago
c 160	14.4	84.7	3762	5	AY816327	AY816327 Homo sapi	c 233	14.4	84.7	159718	4	CNS08CD8	CNS08CD8 Oryza sat
c 161	14.4	84.7	7728	5	AF058907	AF058907 Homo sapi	c 234	14.4	84.7	159749	5	AP003020	AP003020 Oryza sat
c 162	14.4	84.7	9823	15	AE005020	AE005020 Halobacte	c 235	14.4	84.7	160615	4	AC112236	AC112236 Homo sapi
c 163	14.4	84.7	10591	15	AE013950	AE013950 Versinia	c 236	14.4	84.7	161252	12	AC114955	AC114955 Homo sapi
c 164	14.4	84.7	13608	2	AX622895	AX622895 Sequence	c 237	14.4	84.7	162463	6	AC102525	AC102525 Mus muscu

C 238	14.4	84.7	163448	4	OSJN00127	ALG06999 Oryza sat	311	14	82.4	12351	15	AE004035	AE004035 Xylella f
C 239	14.4	84.7	163690	12	AC140908	AC140908 Homo sapi	312	14	82.4	12351	5	AL732409	AL732409 Human DNA
C 240	14.4	84.7	164238	5	AC106745	AC106745 Homo sapi	C 313	14	82.4	20788	15	AE008714	AE008714 Salmonella
C 241	14.4	84.7	165469	5	AC146058	AC146058 Pan trogl	314	14	82.4	37753	15	AF204145	AF204145 Xanthomon
C 242	14.4	84.7	165584	5	AC015922	AC015922 Homo sapi	315	14	82.4	43574	5	AF204145	AF204145 Xanthomon
C 243	14.4	84.7	166871	12	AC170062	AC170062 Sorex ara	316	14	82.4	48561	5	AL162729	AL162729 Homo sapi
C 244	14.4	84.7	167620	5	AC157885	AC157885 Rhinoph	317	14	82.4	50704	12	AC174154	AC174154 Homo sapi
C 245	14.4	84.7	168448	5	AC009469	AC009469 Homo sapi	318	14	82.4	52398	15	CP000157	CP000157 Homo sapi
C 246	14.4	84.7	169194	12	AC025559	AC025559 Homo sapi	C 319	14	82.4	64459	5	CP000157_30	Continuation (31 o
C 247	14.4	84.7	169469	12	AC037195	AC037195 Homo sapi	C 320	14	82.4	64571	6	AL833807	AL833807 Human DNA
C 248	14.4	84.7	169761	12	AC147789	AC147789 Zea mays	C 321	14	82.4	65071	5	AL833807	AL833807 Mouse DNA
C 249	14.4	84.7	171261	12	AC150084	AC150084 Gallus ga	C 322	14	82.4	76561	12	AP001424	AP001424 Homo sapi
C 250	14.4	84.7	171673	12	AC073470	AC073470 Homo sapi	C 323	14	82.4	76561	12	AC121543	AC121543 Mus muscu
C 251	14.4	84.7	173087	5	AC010809	AC010809 Homo sapi	C 324	14	82.4	78291	12	AC023216	AC023216 Homo sapi
C 252	14.4	84.7	174087	12	AC150085	AC150085 Gallus ga	C 325	14	82.4	100000	5	AP000157	AP000157 Homo sapi
C 253	14.4	84.7	174598	5	AC140078	AC140078 Homo sapi	C 326	14	82.4	108400	5	AC005193	AC005193 Homo sapi
C 254	14.4	84.7	174713	12	AC075083	AC075083 Homo sapi	C 327	14	82.4	108971	12	AV914082	AV914082 Triticum
C 255	14.4	84.7	178905	5	AC079316	AC079316 Homo sapi	C 328	14	82.4	110000	4	AE016818_13	Continuation (14 o
C 256	14.4	84.7	181363	11	EX323030	EX323030 Zebratfish	C 329	14	82.4	110000	12	LMFLCHR32_21	Continuation (22 o
C 257	14.4	84.7	181935	5	AC018728	AC018728 Homo sapi	C 330	14	82.4	110000	15	CP000058_01	Continuation (2 of
C 258	14.4	84.7	182788	12	AC173914	AC173914 Nomasus	C 331	14	82.4	110000	15	CP000086_26	Continuation (27 o
C 259	14.4	84.7	184591	6	AC169037	AC169037 Mus muscu	C 332	14	82.4	110000	15	CP000116_27	Continuation (28 o
C 260	14.4	84.7	185415	5	HS121M24	AL354046 Homo sapi	C 333	14	82.4	110000	15	CP000240_03	Continuation (4 of
C 261	14.4	84.7	188869	12	AC177129	AC177129 Strongylo	C 334	14	82.4	110000	15	CR555306_10	Continuation (11 o
C 262	14.4	84.7	189009	12	AC156224	AC156224 Bos tauru	C 335	14	82.4	110000	15	AE009442_14	Continuation (15 o
C 263	14.4	84.7	189479	5	AC027309	AC027309 Homo sapi	C 336	14	82.4	110000	15	AE013598_42	Continuation (43 o
C 264	14.4	84.7	190086	12	AC172086	AC172086 Bos tauru	C 337	14	82.4	110000	15	AE014295_03	Continuation (4 of
C 265	14.4	84.7	191619	12	AC167932	AC167932 Oryctolag	C 338	14	82.4	110000	15	AE014295_13	Continuation (14 o
C 266	14.4	84.7	192375	12	AC022319	AC022319 Homo sapi	C 339	14	82.4	110000	15	AE014613_25	Continuation (26 o
C 267	14.4	84.7	194138	12	AC031270	AC031270 Homo sapi	C 340	14	82.4	110000	15	AE015451_25	Continuation (26 o
C 268	14.4	84.7	194265	5	AC126365	AC126365 Homo sapi	C 341	14	82.4	110000	15	AE017220_04	Continuation (5 of
C 269	14.4	84.7	194291	6	AC087899	AC087899 Mus muscu	C 342	14	82.4	110000	15	AE017220_05	Continuation (6 of
C 270	14.4	84.7	195591	5	AC087499	AC087499 Homo sapi	C 343	14	82.4	110000	15	AP006878_16	Continuation (17 o
C 271	14.4	84.7	199385	12	AC019049	AC019049 Homo sapi	C 344	14	82.4	110000	15	AP008229_42	Continuation (43 o
C 272	14.4	84.7	200384	6	AC130221	AC130221 Mus muscu	C 345	14	82.4	110000	15	AY596227_09	Continuation (10 o
C 273	14.4	84.7	202035	5	AC114291	AC114291 Homo sapi	C 346	14	82.4	110000	15	CP000026_23	Continuation (24 o
C 274	14.4	84.7	204050	15	AL646073	AL646073 Ralstonia	C 347	14	82.4	121310	12	AC141831	AC141831 Apis mell
C 275	14.4	84.7	206380	6	AC007318	AC007318 Homo sapi	C 348	14	82.4	125133	5	AC116631	AC116631 Homo sapi
C 276	14.4	84.7	206380	6	AC113358	AC113358 Mus muscu	C 349	14	82.4	126853	12	AC151508	AC151508 Dasyypu n
C 277	14.4	84.7	208050	15	AJ414144	AJ414144 Yersinia	C 350	14	82.4	129366	5	HS22446	AL031281 Human DNA
C 278	14.4	84.7	212856	6	AL292062	AL292062 Mouse DNA	C 351	14	82.4	138401	12	AC116975	AC116975 Homo sapi
C 279	14.4	84.7	214508	12	AC153263	AC153263 Bos tauru	C 352	14	82.4	139514	12	AC012233	AC012233 Homo sapi
C 280	14.4	84.7	220208	6	AC145748	AC145748 Mus muscu	C 353	14	82.4	141866	5	AL451062	AL451062 Human DNA
C 281	14.4	84.7	223652	12	AC181260	AC181260 Strongylo	C 354	14	82.4	142784	12	AC135544	AC135544 Tetraodon
C 282	14.4	84.7	224591	12	AC159788	AC159788 Bos tauru	C 355	14	82.4	144436	6	AL935331	AL935331 Mouse DNA
C 283	14.4	84.7	231335	12	AC098311	AC098311 Rattus no	C 356	14	82.4	145075	6	AC140515	AC140515 Homo sapi
C 284	14.4	84.7	231771	12	AC178658	AC178658 Strongylo	C 357	14	82.4	145479	6	AC131909	AC131909 Mus muscu
C 285	14.4	84.7	246786	12	AC164802	AC164802 Bos tauru	C 358	14	82.4	148113	12	AL391624	AL391624 Homo sapi
C 286	14.4	84.7	247649	12	AC150040	AC150040 Gallus ga	C 359	14	82.4	150912	12	AC151541	AC151541 Actus nan
C 287	14.4	84.7	257525	12	AC134804	AC134804 Rattus no	C 360	14	82.4	152423	12	AC119145	AC119145 Rattus no
C 288	14.4	84.7	258316	5	AC147975	AC147975 Pan trogl	C 361	14	82.4	155183	12	AC032016	AC032016 Homo sapi
C 289	14.4	84.7	262378	12	AC106581	AC106581 Rattus no	C 362	14	82.4	155683	12	AC105160	AC105160 Mus muscu
C 290	14.4	84.7	276304	12	AC157754	AC157754 Pan trogl	C 363	14	82.4	157572	12	AC127981	AC127981 Rattus no
C 291	14.4	84.7	276304	12	AC157754	AC157754 Pan trogl	C 364	14	82.4	159929	12	AC116973	AC116973 Homo sapi
C 292	14.4	84.7	278920	12	AC171471	AC171471 Bos tauru	C 365	14	82.4	160865	12	AC009818	AC009818 Homo sapi
C 293	14.4	84.7	290924	15	AE017138	AE017138 Yersinia	C 366	14	82.4	162024	6	AL662835	AL662835 Mouse DNA
C 294	14.4	84.7	311000	15	SCO39122	AL939122 Streptomy	C 367	14	82.4	163577	5	AC011848	AC011848 Homo sapi
C 295	14.4	84.7	326002	2	CO870085	CO870085 Sequence	C 368	14	82.4	163674	6	AL669937	AL669937 Mouse DNA
C 296	14.4	84.7	328249	12	AC034210	AC034210 Homo sapi	C 369	14	82.4	165020	5	AC073977	AC073977 Homo sapi
C 297	14.4	84.7	333800	15	SMES91792	AL591792 Slnorthizo	C 370	14	82.4	168744	12	AC174486	AC174486 Strongylo
C 298	14.4	82.4	201	7	BV170733	BV170733 sqm37114	C 371	14	82.4	168931	12	AC144998	AC144998 Pan trogl
C 299	14.4	82.4	338	2	CO449503	AB138001 Homo sapi	C 372	14	82.4	169604	5	AC083884	AC083884 Homo sapi
C 300	14.4	82.4	465	7	AB138001	AB138001 Homo sapi	C 373	14	82.4	169908	12	AC150910	AC150910 Pan trogl
C 301	14.4	82.4	663	4	AF467735	AF467735 Oryza sat	C 374	14	82.4	170809	12	AC141823	AC141823 Apis mell
C 302	14.4	82.4	685	5	BSA331033	AC331033 Homo sapi	C 375	14	82.4	171811	5	AC011739	AC011739 Homo sapi
C 303	14.4	82.4	850	7	BV573069	BV573069 G591P6947	C 376	14	82.4	172963	5	AC134866	AC134866 Homo sapi
C 304	14.4	82.4	1424	15	TFECBBWA	L37437 Thibacillu	C 377	14	82.4	173128	5	AC009127	AC009127 Homo sapi
C 305	14.4	82.4	1439	5	AY336980	AY336980 Homo sapi	C 378	14	82.4	173718	5	AC073575	AC073575 Homo sapi
C 306	14.4	82.4	1439	5	BC047706	BC047706 Homo sapi	C 379	14	82.4	175869	5	AL445489	AL445489 Human DNA
C 307	14.4	82.4	2048	5	AK123716	AK123716 Homo sapi	C 380	14	82.4	176141	5	AC138783	AC138783 Homo sapi
C 308	14.4	82.4	2191	6	BC027807	BC027807 Mus muscu	C 381	14	82.4	179064	6	AC129081	AC129081 Mus muscu
C 309	14.4	82.4	2221	5	AK124334	AK124334 Homo sapi	C 382	14	82.4	182529	5	AC092813	AC092813 Homo sapi
C 310	14.4	82.4	2502	6	AK220237	AK220237 Mus muscu	C 383	14	82.4	182902	5	AC092405	AC092405 Papio anu

C 384	14	82.4	183113	5	HS243E7	AL0232323 Human DNA	457	13.8	81.2	407	10	DQ222885	Human her
C 385	14	82.4	183402	12	AC126231	AC126231 Bos tauru	458	13.8	81.2	407	10	DQ222886	Human her
C 386	14	82.4	188030	12	AC015456	AC015456 Homo sapi	459	13.8	81.2	407	10	DQ223889	Human her
C 387	14	82.4	189174	5	AC092687	AC092687 Homo sapi	460	13.8	81.2	407	10	DQ223890	Human her
C 388	14	82.4	189360	12	AC146696	AC146696 Pan trogl	461	13.8	81.2	407	10	DQ223891	Human her
C 389	14	82.4	189381	12	AC126920	AC126920 Bos tauru	462	13.8	81.2	407	10	DQ223892	Human her
C 390	14	82.4	190200	6	AC114613	AC114613 Mus muscu	463	13.8	81.2	407	10	DQ223893	Human her
C 391	14	82.4	192653	5	AC124781	AC124781 Homo sapi	464	13.8	81.2	407	10	DQ224079	Human her
C 392	14	82.4	195275	6	AC102255	AC102255 Mus muscu	465	13.8	81.2	407	10	DQ224080	Human her
C 393	14	82.4	197719	6	AC168212	AC168212 Mus muscu	466	13.8	81.2	407	10	DQ224081	Human her
C 394	14	82.4	199454	12	AC117180	AC117180 Homo sapi	467	13.8	81.2	407	10	DQ224081	Human her
C 395	14	82.4	198673	6	AL669872	AL669872 Mouse DNA	468	13.8	81.2	407	10	DQ224081	Human her
C 396	14	82.4	199928	12	AC120941	AC120941 Rattus no	469	13.8	81.2	407	10	DQ224080	Human her
C 397	14	82.4	201269	12	AP001849	AP001849 Homo sapi	470	13.8	81.2	407	10	DQ224080	Human her
C 398	14	82.4	201623	12	AL158047	AL158047 Human DNA	471	13.8	81.2	407	10	DQ224081	Human her
C 399	14	82.4	201673	6	AL591544	AL591544 Mouse DNA	472	13.8	81.2	407	10	DQ224081	Human her
C 400	14	82.4	202402	12	AC012517	AC012517 Homo sapi	473	13.8	81.2	407	10	DQ224081	Human her
C 401	14	82.4	204963	12	AC163566	AC163566 Bos tauru	474	13.8	81.2	407	10	DQ224081	Human her
C 402	14	82.4	209121	6	AC127589	AC127589 Mus muscu	475	13.8	81.2	407	10	DQ224081	Human her
C 403	14	82.4	209775	12	AC027219	AC027219 Homo sapi	476	13.8	81.2	407	10	DQ224081	Human her
C 404	14	82.4	211209	12	CT485990	CT485990 Drosophil	477	13.8	81.2	407	10	DQ224081	Human her
C 405	14	82.4	211531	5	AC006995	AC006995 Homo sapi	478	13.8	81.2	407	10	DQ224081	Human her
C 406	14	82.4	213925	12	AC150641	AC150641 Bos tauru	479	13.8	81.2	407	10	DQ224081	Human her
C 407	14	82.4	216211	12	AC091315	AC091315 Mus muscu	480	13.8	81.2	407	10	DQ224081	Human her
C 408	14	82.4	217979	12	AC161613	AC161613 Pan trogl	481	13.8	81.2	407	10	DQ224081	Human her
C 409	14	82.4	220289	6	AC108908	AC108908 Mus muscu	482	13.8	81.2	407	10	DQ224081	Human her
C 410	14	82.4	220932	5	BS000125	BS000125 Pan trogl	483	13.8	81.2	407	10	DQ224081	Human her
C 411	14	82.4	224337	6	AC163434	AC163434 Mus muscu	484	13.8	81.2	407	10	DQ224081	Human her
C 412	14	82.4	228292	6	AL772199	AL772199 Mouse DNA	485	13.8	81.2	407	10	DQ224081	Human her
C 413	14	82.4	229508	5	AC146694	AC146694 Pan trogl	486	13.8	81.2	407	10	DQ224081	Human her
C 414	14	82.4	229898	12	AC161564	AC161564 Bos tauru	487	13.8	81.2	407	10	DQ224081	Human her
C 415	14	82.4	230552	5	AC005098	AC005098 Homo sapi	488	13.8	81.2	407	10	DQ224081	Human her
C 416	14	82.4	232945	12	AC118407	AC118407 Rattus no	489	13.8	81.2	407	10	DQ224081	Human her
C 417	14	82.4	235703	12	AC169692	AC169692 Bos tauru	490	13.8	81.2	407	10	DQ224081	Human her
C 418	14	82.4	239307	12	AC129402	AC129402 Rattus no	491	13.8	81.2	407	10	DQ224081	Human her
C 419	14	82.4	246322	12	AC133966	AC133966 Homo sapi	492	13.8	81.2	407	10	DQ224081	Human her
C 420	14	82.4	246322	12	AC133966	AC133966 Homo sapi	493	13.8	81.2	407	10	DQ224081	Human her
C 421	14	82.4	247376	12	AC156127	AC156127 Bos tauru	494	13.8	81.2	407	10	DQ224081	Human her
C 422	14	82.4	250425	12	AC125667	AC125667 Rattus no	495	13.8	81.2	407	10	DQ224081	Human her
C 423	14	82.4	268050	15	AL627266	AL627266 Salmorell	496	13.8	81.2	407	10	DQ224081	Human her
C 424	14	82.4	275197	5	AC004166	AC004166 Homo sapi	497	13.8	81.2	407	10	DQ224081	Human her
C 425	14	82.4	281138	12	AC159058	AC159058 Bos tauru	498	13.8	81.2	407	10	DQ224081	Human her
C 426	14	82.4	285354	12	AC152597	AC152597 Bos tauru	499	13.8	81.2	407	10	DQ224081	Human her
C 427	14	82.4	292748	12	AC106574	AC106574 Rattus no	500	13.8	81.2	407	10	DQ224081	Human her
C 428	14	82.4	300350	15	BX294145	BX294145 Pirellula	501	13.8	81.2	407	10	DQ224081	Human her
C 429	14	82.4	302521	5	AP000014	AP000014 Homo sapi	502	13.8	81.2	407	10	DQ224081	Human her
C 430	14	82.4	340000	5	AP001730	AP001730 Homo sapi	503	13.8	81.2	407	10	DQ224081	Human her
C 431	14	82.4	349980	2	AX492782	AX492782 Sequence	504	13.8	81.2	407	10	DQ224081	Human her
C 432	14	82.4	349980	2	AX492786	AX492786 Sequence	505	13.8	81.2	407	10	DQ224081	Human her
C 433	14	82.4	349980	2	AX553949	AX553949 Sequence	506	13.8	81.2	407	10	DQ224081	Human her
C 434	14	82.4	349980	2	AX553953	AX553953 Sequence	507	13.8	81.2	407	10	DQ224081	Human her
C 435	13.8	81.2	20	2	AR568092	AR568092 Sequence	508	13.8	81.2	407	10	DQ224081	Human her
C 436	13.8	81.2	47	2	AR290691	AR290691 Sequence	509	13.8	81.2	407	10	DQ224081	Human her
C 437	13.8	81.2	117	15	NPE564413	AJ564413 Mycobacte	510	13.8	81.2	407	10	DQ224081	Human her
C 438	13.8	81.2	117	15	MXE564419	AJ564419 Mycobacte	511	13.8	81.2	407	10	DQ224081	Human her
C 439	13.8	81.2	123	5	AF439454	AF439454 Homo sapi	512	13.8	81.2	407	10	DQ224081	Human her
C 440	13.8	81.2	201	7	BV171711	BV171711 sqm45306	513	13.8	81.2	407	10	DQ224081	Human her
C 441	13.8	81.2	206	7	CR382032	CR382032 Arabidops	514	13.8	81.2	407	10	DQ224081	Human her
C 442	13.8	81.2	209	6	WGHIF3A06	AF079145 Mus muscu	515	13.8	81.2	407	10	DQ224081	Human her
C 443	13.8	81.2	219	10	AF134739	AF134739 Hepatitis	516	13.8	81.2	407	10	DQ224081	Human her
C 444	13.8	81.2	228	2	AC131360	AC131360 Sequence	517	13.8	81.2	407	10	DQ224081	Human her
C 445	13.8	81.2	231	5	HS13F5R	Z54608 H.sapiens C	518	13.8	81.2	407	10	DQ224081	Human her
C 446	13.8	81.2	240	10	HS1LRNA1	K03352 HSV1 (KOS) ,	519	13.8	81.2	407	10	DQ224081	Human her
C 447	13.8	81.2	247	8	SYNLRHBCA	M22006 Synthetic h	520	13.8	81.2	407	10	DQ224081	Human her
C 448	13.8	81.2	273	7	BV0000760	BV0000760 S210P6198	521	13.8	81.2	407	10	DQ224081	Human her
C 449	13.8	81.2	283	2	AR247212	AR247212 Sequence	522	13.8	81.2	407	10	DQ224081	Human her
C 450	13.8	81.2	293	2	AR249750	AR249750 Sequence	523	13.8	81.2	407	10	DQ224081	Human her
C 451	13.8	81.2	320	13	AY465871	AY465871 Clinostom	524	13.8	81.2	407	10	DQ224081	Human her
C 452	13.8	81.2	350	2	CQ458356	CQ458356 Sequence	525	13.8	81.2	407	10	DQ224081	Human her
C 453	13.8	81.2	363	2	AX308452	AX308452 Sequence	526	13.8	81.2	407	10	DQ224081	Human her
C 454	13.8	81.2	375	15	AV726000	AV726000 Pseudomon	527	13.8	81.2	407	10	DQ224081	Human her
C 455	13.8	81.2	385	2	AR149028	AR149028 Sequence	528	13.8	81.2	407	10	DQ224081	Human her
C 456	13.8	81.2	385	2	AR340400	AR340400 Sequence	529	13.8	81.2	407	10	DQ224081	Human her
C 457	13.8	81.2	385	2	AR340400	AR340400 Sequence	530	13.8	81.2	407	10	DQ224081	Human her
C 458	13.8	81.2	385	2	AR340400	AR340400 Sequence	531	13.8	81.2	407	10	DQ224081	Human her
C 459	13.8	81.2	385	2	AR340400	AR340400 Sequence	532	13.8	81.2	407	10	DQ224081	Human her
C 460	13.8	81.2	385	2	AR340400	AR340400 Sequence	533	13.8	81.2	407	10	DQ224081	Human her
C 461	13.8	81.2	385	2	AR340400	AR340400 Sequence	534	13.8	81.2	407	10	DQ224081	Human her
C 462	13.8	81.2	385	2	AR340400	AR340400 Sequence	535	13.8	81.2	407	10	DQ224081	Human her
C 463	13.8	81.2	385	2	AR340400	AR340400 Sequence	536	13.8	81.2	407	10	DQ224081	Human her
C 464	13.8	81.2	385	2	AR340400	AR340400 Sequence	537	13.8	81.2	407	10	DQ224081	Human her
C 465	13.8	81.2	385	2	AR340400	AR340400 Sequence	538	13.8	81.2	407	10	DQ224081	Human her
C 466	13.8	81.2	385	2	AR340400	AR340400 Sequence	539	13.8	81.2	407	10	DQ224081	Human her
C 467	13.8	81.2	385	2	AR340400	AR340400 Sequence	540	13.8	81.2	407	10	DQ224081	Human her
C 468	13.8	81.2	385	2	AR340400	AR340400 Sequence	541	13.8	81.2	407	10	DQ224081	Human her
C 469	13.8	81.2	385	2	AR340400	AR340400 Sequence	542	13.8	81.2	407	10	DQ224081	Human her
C 470	13.8	81.2	385	2	AR340400	AR340400 Sequence	543	13.8	81.2	407	10	DQ224081	Human her
C 471	13.8	81.2	385	2	AR340400	AR340400 Sequence	544	13.8	81.2	407	10	DQ224081	Human her
C 472	13.8	81.2	385	2	AR340400	AR340400 Sequence	545	13.8	81.2	407	10	DQ224081	Human her
C 473	13.8	81.2	385	2	AR340400	AR340400 Sequence	546	13.8	81.2	407	10	DQ224081	Human her
C 474	13.8	81.2	385	2	AR340400	AR340400 Sequence	547	13.8	81.2	407	10	DQ224081	Human her
C 475	13.8	81.2	385	2	AR340400	AR340400 Sequence	548	13.8	81.2	407	10	DQ224081	Human her
C 476	13.8	81.2	385	2	AR340400	AR340400 Sequence							

530	13.8	81.2	781	7	BV044236	BV044236	S212P6023	C 603	13.8	81.2	1263	15	AB014207	AB014207 Mycobacte
531	13.8	81.2	791	7	BV564649	BV564649	QZG60D03	C 604	13.8	81.2	1305	2	CQ715205	CQ715205 Sequence
C 532	13.8	81.2	793	7	BV589324	BV589324	G591P6354	C 605	13.8	81.2	1317	2	BD180606	BD180606 Sequence
C 533	13.8	81.2	795	2	ARE24063	ARE24063	G591P6354	C 606	13.8	81.2	1335	2	AX653009	AX653009 Sequence
C 534	13.8	81.2	799	2	BD018067	BD018067	Novel gen	607	13.8	81.2	1350	14	AB211528	AB211528 Canis fam
C 535	13.8	81.2	799	2	BD098005	BD098005	Novel gen	608	13.8	81.2	1410	14	S83090	S83090 cholecystok
536	13.8	81.2	806	7	BV487824	BV487824	S212P6214	609	13.8	81.2	1420	6	BC096647	BC096647 Mus muscu
537	13.8	81.2	808	7	BV590297	BV590297	G591P6427	610	13.8	81.2	1475	15	MKHM0G	MKHM0G Mus muscu
C 538	13.8	81.2	842	7	BV557155	BV557155	QJ809G10	611	13.8	81.2	1508	11	EX935103	EX935103 Gallus ga
C 539	13.8	81.2	857	6	AV686592	AV686592	Rattus no	612	13.8	81.2	1521	14	DQ288287	DQ288287 Ovis arie
C 540	13.8	81.2	861	2	A12338	A12338	tfdA-2,4-D-	613	13.8	81.2	1536	14	DQ288286	DQ288286 Bos tauru
C 541	13.8	81.2	862	6	MMILGFB4	MMILGFB4	Musculus	614	13.8	81.2	1546	11	AY138245	AY138245 Gallus ga
C 542	13.8	81.2	904	7	BV576772	BV576772	G591P6094	C 615	13.8	81.2	1548	4	AY342321	AY342321 Oryza sat
C 543	13.8	81.2	906	15	STMPAC	M25346	Streptomyce	C 616	13.8	81.2	1578	2	A38265	A38265 Sequence 1
C 544	13.8	81.2	915	13	AF466828	AF466828	Asterina	C 617	13.8	81.2	1578	2	AR001081	AR001081 Sequence
C 545	13.8	81.2	946	11	EX931799	EX931799	Gallus ga	C 618	13.8	81.2	1604	5	AR007549	AR007549 Homo sapi
546	13.8	81.2	951	2	CS070401	CS070401	Sequence	C 619	13.8	81.2	1616	2	AR581370	AR581370 Sequence
547	13.8	81.2	958	13	AF470547	AF470547	Bolbophor	C 620	13.8	81.2	1635	5	HUMBHEXB	M19735 Homo sapien
C 548	13.8	81.2	960	7	BV562186	BV562186	soc97f09	C 621	13.8	81.2	1653	4	LSU18510	U18510 Leptosira s
C 549	13.8	81.2	965	4	AK063675	AK063675	Oryza sat	C 622	13.8	81.2	1669	5	AB035747	AB035747 Homo sapi
550	13.8	81.2	975	5	HSY16275	Y16275	Homo sapien	C 623	13.8	81.2	1671	5	BT009919	BT009919 Homo sapi
551	13.8	81.2	1010	11	EX933364	EX933364	Gallus ga	C 624	13.8	81.2	1671	8	AY890181	AY890181 Synthetic
552	13.8	81.2	1023	13	AF470541	AF470541	Bolbophor	C 625	13.8	81.2	1671	8	AY892668	AY892668 Synthetic
553	13.8	81.2	1023	13	AF470544	AF470544	Bolbophor	C 626	13.8	81.2	1671	8	BT007907	BT007907 Synthetic
554	13.8	81.2	1023	13	AF470572	AF470572	Bolbophor	C 627	13.8	81.2	1672	2	CS108503	CS108503 Sequence
555	13.8	81.2	1023	13	AF470575	AF470575	Bolbophor	C 628	13.8	81.2	1678	15	CCU13664	U13664 Caulobacter
556	13.8	81.2	1023	13	AF470579	AF470579	Bolbophor	C 629	13.8	81.2	1680	2	AX078609	AX078609 Sequence
557	13.8	81.2	1023	13	AF470583	AF470583	Bolbophor	C 630	13.8	81.2	1682	4	AF387152	AF387152 Geminella
558	13.8	81.2	1023	13	AF470587	AF470587	Bolbophor	C 631	13.8	81.2	1686	4	DQ009743	DQ009743 Asteromon
559	13.8	81.2	1023	13	AF470591	AF470591	Bolbophor	C 632	13.8	81.2	1687	4	DQ009744	DQ009744 Asteromon
560	13.8	81.2	1023	13	AF470595	AF470595	Bolbophor	C 633	13.8	81.2	1688	4	AY220580	AY220580 Chlamydom
561	13.8	81.2	1023	13	AF470599	AF470599	Bolbophor	C 634	13.8	81.2	1695	5	AF242529	AF242529 Homo sapi
562	13.8	81.2	1023	13	AF470603	AF470603	Bolbophor	C 635	13.8	81.2	1739	4	AB078732	AB078732 Wittrocki
563	13.8	81.2	1023	13	AF470607	AF470607	Bolbophor	C 636	13.8	81.2	1741	4	AB062717	AB062717 Wittrocki
564	13.8	81.2	1023	13	AF470610	AF470610	Bolbophor	C 637	13.8	81.2	1755	4	AB058305	AB058305 Chlorella
565	13.8	81.2	1024	13	AF470553	AF470553	Bolbophor	C 638	13.8	81.2	1782	4	AK070371	AK070371 Oryza sat
566	13.8	81.2	1024	13	AF470556	AF470556	Bolbophor	C 639	13.8	81.2	1783	4	ASORGEA	M95614 Asteromonas
567	13.8	81.2	1024	13	AF470560	AF470560	Bolbophor	C 640	13.8	81.2	1784	2	BD187623	BD187623 Screening
568	13.8	81.2	1024	13	AF470564	AF470564	Bolbophor	C 641	13.8	81.2	1792	1	ECL402355	EAJ02355 eukaryote
569	13.8	81.2	1024	13	AF470568	AF470568	Bolbophor	C 642	13.8	81.2	1810	5	AK130375	AK130375 Homo sapi
C 570	13.8	81.2	1024	13	AF470568	AF470568	Bolbophor	C 643	13.8	81.2	1824	5	BC017378	BC017378 Homo sapi
C 571	13.8	81.2	1065	2	CQ719536	CQ719536	Sequence	C 644	13.8	81.2	1832	5	HUMBHAB	M13519 Human N-ace
C 572	13.8	81.2	1071	10	S83529	S83529	Bovine leuk	C 645	13.8	81.2	1841	2	CQ727216	CQ727216 Sequence
C 573	13.8	81.2	1090	5	HUMHEXB01	M23282	Human beta-	C 646	13.8	81.2	1881	5	AK130002	AK130002 Homo sapi
C 574	13.8	81.2	1111	2	AX234332	AX234332	Sequence	C 647	13.8	81.2	1892	5	AF378118	AF378118 Homo sapi
C 575	13.8	81.2	1122	4	AF061209	AF061209	Oryza sat	C 648	13.8	81.2	1917	11	AY138247	AY138247 Gallus ga
576	13.8	81.2	1147	5	AF048995	AF048995	Homo sapi	C 649	13.8	81.2	1930	6	MMIGFBP4	Z95492 M.musculus
577	13.8	81.2	1150	4	BD223780	BD223780	Human RNA	C 650	13.8	81.2	1958	5	AF124143	AF124143 Homo sapi
578	13.8	81.2	1158	4	SFDNMTA1	Y15168	Sordaria fi	C 651	13.8	81.2	1976	5	AF059202	AF059202 Homo sapi
579	13.8	81.2	1158	5	BC002973	BC002973	Homo sapi	C 652	13.8	81.2	1980	4	AK099306	AK099306 Oryza sat
580	13.8	81.2	1161	2	AX081385	AX081385	Sequence	C 653	13.8	81.2	1986	11	AF178670	AF178670 Gallus ga
C 581	13.8	81.2	1167	15	AB014917	AB014917	Chromatiu	C 654	13.8	81.2	1996	4	AK069202	AK069202 Oryza sat
C 582	13.8	81.2	1168	2	CQ981341	CQ981341	Sequence	C 655	13.8	81.2	1998	5	BC015762	BC015762 Homo sapi
583	13.8	81.2	1168	5	AF039652	AF039652	Homo sapi	C 656	13.8	81.2	2000	2	AX656916	AX656916 Sequence
C 584	13.8	81.2	1176	6	MMILGFBP4	X81582	M.musculus	C 657	13.8	81.2	2000	5	BC023565	BC023565 Homo sapi
C 585	13.8	81.2	1176	15	AB014153	AB014153	Micromono	C 658	13.8	81.2	2012	10	AV751737	AV751737 Vicia cry
C 586	13.8	81.2	1179	15	AB014145	AB014145	Micromono	C 659	13.8	81.2	2044	5	AK130707	AK130707 Homo sapi
C 587	13.8	81.2	1179	15	AB014156	AB014156	Micromono	C 660	13.8	81.2	2058	2	A12337	A12337 tfda gene c
C 588	13.8	81.2	1179	15	AB014161	AB014161	Micromono	C 661	13.8	81.2	2058	15	AFATFPA	M16730 A.eutrophus
C 589	13.8	81.2	1179	15	AB014163	AB014163	Micromono	C 662	13.8	81.2	2079	6	BC019836	BC019836 Mus muscu
C 590	13.8	81.2	1179	15	AB015622	AB015622	Micromono	C 663	13.8	81.2	2092	5	BD130509	BD130509 Secretary
591	13.8	81.2	1182	2	AR449602	AR449602	Sequence	C 664	13.8	81.2	2092	5	AF250393	AF250393 Homo sapi
C 592	13.8	81.2	1184	15	AB014891	AB014891	Burkholde	C 665	13.8	81.2	2117	5	BC016739	BC016739 Homo sapi
C 593	13.8	81.2	1194	15	AB014944	AB014944	Rhodobact	C 666	13.8	81.2	2120	2	AX702622	AX702622 Sequence
C 594	13.8	81.2	1197	13	AY2568145	AY2568145	Leucochlo	C 667	13.8	81.2	2122	6	BC017641	BC017641 Mus muscu
C 595	13.8	81.2	1200	5	AF262988	AF262988	Homo sapi	C 668	13.8	81.2	2123	5	BC004465	BC004465 Homo sapi
C 596	13.8	81.2	1228	11	CR523425	CR523425	Gallus ga	C 669	13.8	81.2	2123	5	BC005841	BC005841 Homo sapi
C 597	13.8	81.2	1235	2	I72656	I72656	Sequence 13	C 670	13.8	81.2	2124	4	ATH507211	ATH507211 Arabidops
598	13.8	81.2	1235	2	I72657	I72657	Sequence 15	C 671	13.8	81.2	2129	5	BC078171	BC078171 Homo sapi
C 599	13.8	81.2	1239	13	AB062237	AB062237	Leucochlo	C 672	13.8	81.2	2140	5	BC022428	BC022428 Homo sapi
C 600	13.8	81.2	1263	2	BD062237	BD062237	Nucleic a	C 673	13.8	81.2	2146	5	AK000669	AK000669 Homo sapi
C 601	13.8	81.2	1263	2	E44174	E44174	Methods of	C 674	13.8	81.2	2165	5	AK000669	AK000669 Homo sapi
C 602	13.8	81.2	1263	2	AX137458	AX137458	Sequence	C 675	13.8	81.2	2165	5	AK000669	AK000669 Homo sapi

c 676	13.8	81.2	2165	6	BC090335	BC090335 Rattus no	c 749	13.8	81.2	4317	5	BC040228	BC040228 Homo sapi
c 677	13.8	81.2	2169	2	AR447719	AR447719 Sequence	c 750	13.8	81.2	4319	2	AX405634	AX405634 Sequence
678	13.8	81.2	2174	2	AR225609	AR225609 Sequence	752	13.8	81.2	4364	8	AY497507	AY497507 Cloning v
679	13.8	81.2	2174	2	AR692065	AR692065 Sequence	753	13.8	81.2	4381	2	DD165226	DD165226 NUCLEIC A
680	13.8	81.2	2174	2	AX365106	AX365106 Sequence	754	13.8	81.2	4395	11	DQ173578	DQ173578 Pan trogl
681	13.8	81.2	2184	2	AR623715	AR623715 Sequence	755	13.8	81.2	4455	11	XLU14164	XLU14164 Xenopus lae
c 682	13.8	81.2	2194	6	BC098750	BC098750 Rattus no	755	13.8	81.2	4561	8	AY745747	AY745747 pGeneClip
c 683	13.8	81.2	2205	6	BC012270	BC012270 Mus muscu	756	13.8	81.2	4563	5	BC041767	BC041767 Homo sapi
c 684	13.8	81.2	2207	6	AF060194	AF060194 Mus muscu	c 757	13.8	81.2	4579	15	PSEOPRC	D28119 Pseudomonas
c 685	13.8	81.2	2318	15	AF291753	AF291753 Pseudomon	758	13.8	81.2	4611	15	RSRISAFEG	X87256 R.sphaeroid
c 686	13.8	81.2	2328	15	PSEALG76	M37181 P.aeruginos	c 759	13.8	81.2	4617	15	PAU27829	U27829 Pseudomonas
687	13.8	81.2	2351	2	CS113293	CS113293 Sequence	c 760	13.8	81.2	4625	5	AF605898	AF605898 Homo sapi
688	13.8	81.2	2357	10	AY363172	AY363172 Pseudorab	761	13.8	81.2	4639	8	AF346623	AF346623 RAGE vect
c 689	13.8	81.2	2417	2	CQ604580	CQ604580 Sequence	762	13.8	81.2	4783	2	CQ850207	CQ850207 Sequence
c 690	13.8	81.2	2433	6	AB041557	AB041557 Mus muscu	763	13.8	81.2	4783	5	AK128774	AK128774 Homo sapi
c 691	13.8	81.2	2434	6	AB041559	AB041559 Mus muscu	764	13.8	81.2	4852	2	CQ584930	CQ584930 Sequence
692	13.8	81.2	2477	14	BC112800	BC112800 Bos Tauru	765	13.8	81.2	4951	8	ASAJ154	AY000154 Artificia
693	13.8	81.2	2501	2	CQ778570	CQ778570 Sequence	766	13.8	81.2	5043	2	CS165130	CS165130 Sequence
c 694	13.8	81.2	2510	10	SH1UL2XA	M95285 Suid herpes	767	13.8	81.2	5049	2	CQ598359	CQ598359 Sequence
c 695	13.8	81.2	2557	4	AY091763	AY091763 Arabidops	768	13.8	81.2	5049	2	CQ847818	CQ847818 Sequence
696	13.8	81.2	2573	2	CQ843419	CQ843419 Sequence	769	13.8	81.2	5101	2	CS165129	CS165129 Sequence
697	13.8	81.2	2573	5	AK126380	AK126380 Homo sapi	770	13.8	81.2	5139	15	AY729020	AY729020 Rhizobium
698	13.8	81.2	2585	4	AK106303	AK106303 Oryza sat	771	13.8	81.2	5177	6	AF071086	AF071086 Mus muscu
699	13.8	81.2	2626	6	AB041588	AB041588 Mus muscu	c 772	13.8	81.2	5204	2	CQ818774	CQ818774 Sequence
700	13.8	81.2	2634	5	HSAS4981	AY564981 Homo sapi	c 773	13.8	81.2	5235	11	AB081299	AB081299 Sililago j
701	13.8	81.2	2787	2	CQ584931	CQ584931 Sequence	774	13.8	81.2	5325	8	ASPREESP	ZV5185 Artificial
702	13.8	81.2	2893	13	AY245701	AY245701 Clinostom	775	13.8	81.2	5351	8	AY024358	AY024358 Integrati
703	13.8	81.2	2893	13	BT023853	BT023853 Drosophil	776	13.8	81.2	5382	2	AR215117	AR215117 Sequence
704	13.8	81.2	2916	15	PME012480	AY012480 Pseudomon	777	13.8	81.2	5382	2	AR302359	AR302359 Sequence
c 705	13.8	81.2	2925	5	HUMCA1IA01	M77176 H.sapiens c	778	13.8	81.2	5382	2	AR373231	AR373231 Sequence
706	13.8	81.2	2974	13	AY242851	AY242851 Bolbophor	779	13.8	81.2	5382	2	AR401617	AR401617 Sequence
707	13.8	81.2	3027	13	AY122082	AY122082 Drosophil	780	13.8	81.2	5382	2	AR442950	AR442950 Sequence
708	13.8	81.2	3110	6	BC013092	BC013092 Mus muscu	781	13.8	81.2	5382	2	AR568592	AR568592 Sequence
709	13.8	81.2	3133	2	I15000	I15000 Sequence 1	782	13.8	81.2	5382	2	AR658613	AR658613 Sequence
710	13.8	81.2	3135	11	GGCKR6A	Z19110 G.gallus Ce	c 783	13.8	81.2	5382	15	AB003906	AB003906 Hydrogeno
711	13.8	81.2	3179	2	CQ849853	CQ849853 Sequence	784	13.8	81.2	5466	5	AB209638	AB209638 Homo sapi
712	13.8	81.2	3179	5	AK128915	AK128915 Homo sapi	c 785	13.8	81.2	5618	15	AY622309	AY622309 Pectobact
713	13.8	81.2	3216	6	BC085251	BC085251 Mus muscu	786	13.8	81.2	5630	5	AY375845	AY375845 Homo sapi
714	13.8	81.2	3297	8	AF090453	AF090453 Cloning v	787	13.8	81.2	5648	5	AY375847	AY375847 Homo sapi
715	13.8	81.2	3357	2	CQ721036	CQ721036 Sequence	788	13.8	81.2	5650	8	DQ172901	DQ172901 Transfect
c 716	13.8	81.2	3360	2	CQ591030	CQ591030 Sequence	789	13.8	81.2	5656	8	CR847878	CR847878 Condition
c 717	13.8	81.2	3387	11	AF508797	AB231587 Danio rer	790	13.8	81.2	5664	5	AY375851	AY375851 Homo sapi
c 718	13.8	81.2	3438	11	AF508797	AF508797 Gallus ga	791	13.8	81.2	5668	5	AY375852	AY375852 Homo sapi
c 719	13.8	81.2	3453	2	AR623312	AR623312 Sequence	792	13.8	81.2	5693	5	AY375854	AY375854 Homo sapi
720	13.8	81.2	3500	5	AF442769	AF442769 Homo sapi	793	13.8	81.2	5697	5	AY375854	AY375854 Homo sapi
721	13.8	81.2	3544	4	AB020748	AB020748 Arabidops	794	13.8	81.2	5699	5	AY375848	AY375848 Homo sapi
722	13.8	81.2	3558	5	HUMDRBU	M15809 Human MHC C	795	13.8	81.2	5703	5	AY375849	AY375849 Homo sapi
723	13.8	81.2	3575	5	HSHFXH4	X94553 H.sapiens H	796	13.8	81.2	5706	8	AF529183	AF529183 Episomal
c 724	13.8	81.2	3604	6	AF348523	AF348523 Rattus no	797	13.8	81.2	5707	5	AY375853	AY375853 Homo sapi
725	13.8	81.2	3636	2	AX714273	AX714273 Sequence	798	13.8	81.2	5711	5	AY375850	AY375850 Homo sapi
c 726	13.8	81.2	3636	5	AK056495	AK056495 Homo sapi	c 799	13.8	81.2	5936	8	AY613996	AY613996 Cloning v
c 727	13.8	81.2	3657	6	AY749168	AY749168 Mus muscu	c 800	13.8	81.2	6066	2	CQ591029	CQ591029 Sequence
c 728	13.8	81.2	3668	15	RTNODG	X03721 Rhizobium t	c 801	13.8	81.2	6076	8	AB086386	AB086386 Retrovira
c 729	13.8	81.2	3674	5	AF198358	AF198358 Homo sapi	c 802	13.8	81.2	6113	15	AF393183	AF393183 Clavibact
c 730	13.8	81.2	3729	2	CQ850157	CQ850157 Sequence	803	13.8	81.2	6233	2	CS165135	CS165135 Sequence
c 731	13.8	81.2	3729	5	AK127261	AK127261 Homo sapi	804	13.8	81.2	6279	2	AX207276	AX207276 Sequence
c 732	13.8	81.2	3747	6	AF013131	AF013131 Mus muscu	805	13.8	81.2	6289	5	AB018339	AB018339 Homo sapi
c 733	13.8	81.2	3752	11	FRD1	X80174 F.rubripes	c 806	13.8	81.2	6381	2	BD171160	BD171160 Novel gen
c 734	13.8	81.2	3766	13	BT011380	BT011380 Drosophil	c 807	13.8	81.2	6381	2	BD183438	BD183438 Novel gen
c 735	13.8	81.2	3797	10	PVULSGENE	X87247 Pseudorabie	c 808	13.8	81.2	6381	5	AB051439	AB051439 Homo sapi
c 736	13.8	81.2	3814	2	AR531687	AR531687 Sequence	c 809	13.8	81.2	6381	15	MKHMMDTD	Y10251 M.kandleri
c 737	13.8	81.2	3956	2	AX801721	AX801721 Sequence	c 810	13.8	81.2	6515	2	CS175841	CS175841 Sequence
c 738	13.8	81.2	3956	6	AF074266	AF074266 Mus muscu	c 811	13.8	81.2	6515	2	CS085795	CS085795 Sequence
c 739	13.8	81.2	4040	8	AY613994	AY613994 Cloning v	812	13.8	81.2	6561	2	CS085795	CS085795 Sequence
c 740	13.8	81.2	4071	6	AB011439	AB011439 Rattus no	813	13.8	81.2	6630	13	AY197345	AY197345 Schistosom
c 741	13.8	81.2	4133	10	AB218901	AB218901 Chimpanze	814	13.8	81.2	6637	13	AY197343	AY197343 Schistosom
742	13.8	81.2	4173	8	AY508732	AY508732 Cloning v	815	13.8	81.2	6657	13	AF197344	AF197344 Schistosom
743	13.8	81.2	4192	5	BC043183	BC043183 Homo sapi	816	13.8	81.2	6762	8	AF334709	AF334709 Envelope
c 744	13.8	81.2	4214	5	AB037800	AB037800 Homo sapi	817	13.8	81.2	7122	2	CS175842	CS175842 Sequence
745	13.8	81.2	4257	2	CQ918571	CQ918571 Sequence	c 818	13.8	81.2	7180	2	AX468870	AX468870 Sequence
746	13.8	81.2	4257	8	CVU07648	CVU07648 Cloning vec	819	13.8	81.2	7201	2	CQ889769	CQ889769 Sequence
747	13.8	81.2	4282	5	DQ173577	DQ173577 Homo sapi	c 820	13.8	81.2	7201	2	CQ973535	CQ973535 Sequence
c 748	13.8	81.2	4307	11	BC079756	BC079756 Xenopus 1	c 821	13.8	81.2	7201	2	AX800664	AX800664 Sequence

C 822	13.8	81.2	7231	2	AX806466	AX806466 Sequence	C 895	13.8	81.2	10637	5	AY184206	AY184206 Homo sapi
C 823	13.8	81.2	7291	8	ASPSVPAZ1	Y07573 pSPa211 pl	C 896	13.8	81.2	10654	8	CR293541	CR293541 3'HPT (M
C 824	13.8	81.2	7294	8	ASPSVPAZ1	Y07631 pSPaX1 pla	C 897	13.8	81.2	10742	2	CS036440	CS036440 Sequence
C 825	13.8	81.2	7297	2	AX806467	AX806467 Sequence	C 898	13.8	81.2	10742	2	CS036440	CS036440 Sequence
C 826	13.8	81.2	7353	4	CNS09540	AX806467 Sequence	C 899	13.8	81.2	10742	5	AY061755	AY061755 Homo sapi
C 827	13.8	81.2	7429	2	AX839727	AX839727 Sequence	C 900	13.8	81.2	10774	2	CQ789660	CQ789660 Sequence
C 828	13.8	81.2	7438	2	AX816382	AX816382 Sequence	C 901	13.8	81.2	10805	12	AC017707	AC017707 Drosophila
C 829	13.8	81.2	7456	2	AX806465	AX806465 Sequence	C 902	13.8	81.2	10868	8	DQ322637	DQ322637 VEEV repl
C 830	13.8	81.2	7504	5	AB040943	AB040943 Homo sapi	C 903	13.8	81.2	10921	2	CQ789658	CQ789658 Sequence
C 831	13.8	81.2	7573	2	AX114856	AX114856 Sequence	C 904	13.8	81.2	10928	8	DQ322636	DQ322636 VEEV repl
C 832	13.8	81.2	7573	2	AX114856	AX114856 Sequence	C 905	13.8	81.2	10928	8	DQ322636	DQ322636 VEEV repl
C 833	13.8	81.2	7800	10	HEHSV1EX	AX114870 Sequence	C 906	13.8	81.2	10928	8	DQ322636	DQ322636 VEEV repl
C 834	13.8	81.2	8050	15	AE006033	Y07573 pSPa211 pl	C 907	13.8	81.2	10928	8	DQ322636	DQ322636 VEEV repl
C 835	13.8	81.2	8157	12	AL831743	AX08339 HSV-1 (stra	C 908	13.8	81.2	10959	15	AE004953	AE004953 Pseudomon
C 836	13.8	81.2	8298	2	AL831743	AX08339 HSV-1 (stra	C 909	13.8	81.2	10961	2	CQ789656	CQ789656 Sequence
C 837	13.8	81.2	8454	2	CNS09540	AX08339 HSV-1 (stra	C 910	13.8	81.2	10978	15	AE012135	AE012135 Xanthomon
C 838	13.8	81.2	8540	2	AF198722	AX08339 HSV-1 (stra	C 911	13.8	81.2	11006	2	CQ789654	CQ789654 Sequence
C 839	13.8	81.2	8588	10	AF257515	AX08339 HSV-1 (stra	C 912	13.8	81.2	11059	2	CQ789683	CQ789683 Sequence
C 840	13.8	81.2	8742	2	AX553924	AX08339 HSV-1 (stra	C 913	13.8	81.2	11242	15	AE011933	AE011933 Xanthomon
C 841	13.8	81.2	8815	4	SMSMTA	AX08339 HSV-1 (stra	C 914	13.8	81.2	11407	15	AE004775	AE004775 Pseudomon
C 842	13.8	81.2	9027	2	AX577791	AX08339 HSV-1 (stra	C 915	13.8	81.2	11646	2	AX577790	AX577790 Sequence
C 843	13.8	81.2	9320	2	AX663075	AX08339 HSV-1 (stra	C 916	13.8	81.2	12041	2	AX577789	AX577789 Sequence
C 844	13.8	81.2	9382	10	SHIULGNS	AX08339 HSV-1 (stra	C 917	13.8	81.2	12055	15	AE009467	AE009467 Brucella
C 845	13.8	81.2	9482	2	CQ829527	AX08339 HSV-1 (stra	C 918	13.8	81.2	12292	15	AE009033	AE009033 Agrobacte
C 846	13.8	81.2	9482	2	CQ829527	AX08339 HSV-1 (stra	C 919	13.8	81.2	12378	15	AE004797	AE004797 Pseudomon
C 847	13.8	81.2	9737	2	AR215118	AX08339 HSV-1 (stra	C 920	13.8	81.2	12575	15	AE007999	AE007999 Agrobacte
C 848	13.8	81.2	9737	2	AR215118	AX08339 HSV-1 (stra	C 921	13.8	81.2	12753	15	EBWCRS	EBWCRS Pantoea agg
C 849	13.8	81.2	9737	2	AR215119	AX08339 HSV-1 (stra	C 922	13.8	81.2	12840	8	DQ322638	DQ322638 VEEV repl
C 850	13.8	81.2	9737	2	AR302360	AX08339 HSV-1 (stra	C 923	13.8	81.2	12840	8	DQ322640	DQ322640 VEEV repl
C 851	13.8	81.2	9737	2	AR302361	AX08339 HSV-1 (stra	C 924	13.8	81.2	12840	8	DQ322642	DQ322642 VEEV repl
C 852	13.8	81.2	9737	2	AR302366	AX08339 HSV-1 (stra	C 925	13.8	81.2	13068	2	CQ790429	CQ790429 Sequence
C 853	13.8	81.2	9737	2	AR373232	AX08339 HSV-1 (stra	C 926	13.8	81.2	13547	2	AX577782	AX577782 Sequence
C 854	13.8	81.2	9737	2	AR373232	AX08339 HSV-1 (stra	C 927	13.8	81.2	13766	2	AX553919	AX553919 Sequence
C 855	13.8	81.2	9737	2	AR373238	AX08339 HSV-1 (stra	C 928	13.8	81.2	14136	11	TRU544612	TRU544612 Takifugu
C 856	13.8	81.2	9737	2	AR401618	AX08339 HSV-1 (stra	C 929	13.8	81.2	14338	15	AE010303	AE010303 Myxococc
C 857	13.8	81.2	9737	2	AR401619	AX08339 HSV-1 (stra	C 930	13.8	81.2	14863	8	AY339820	AY339820 Nx-Lox ta
C 858	13.8	81.2	9737	2	AR401624	AX08339 HSV-1 (stra	C 931	13.8	81.2	14987	2	CS058097	CS058097 Sequence
C 859	13.8	81.2	9737	2	AR442951	AX08339 HSV-1 (stra	C 932	13.8	81.2	15081	2	CQ790448	CQ790448 Sequence
C 860	13.8	81.2	9737	2	AR442952	AX08339 HSV-1 (stra	C 933	13.8	81.2	15083	2	CS101424	CS101424 Sequence
C 861	13.8	81.2	9737	2	AR442957	AX08339 HSV-1 (stra	C 934	13.8	81.2	15110	2	CS101426	CS101426 Sequence
C 862	13.8	81.2	9737	2	AR568593	AX08339 HSV-1 (stra	C 935	13.8	81.2	15407	11	AF536192	AF536192 Gallus ga
C 863	13.8	81.2	9737	2	AR568594	AX08339 HSV-1 (stra	C 936	13.8	81.2	17290	2	AX708622	AX708622 Sequence
C 864	13.8	81.2	9737	2	AR568599	AX08339 HSV-1 (stra	C 937	13.8	81.2	17344	2	CQ599660	CQ599660 Sequence
C 865	13.8	81.2	9737	2	AR568614	AX08339 HSV-1 (stra	C 938	13.8	81.2	17358	2	CQ584942	CQ584942 Sequence
C 866	13.8	81.2	9737	2	AR568615	AX08339 HSV-1 (stra	C 939	13.8	81.2	17402	2	CS165133	CS165133 Sequence
C 867	13.8	81.2	9737	2	AR568620	AX08339 HSV-1 (stra	C 940	13.8	81.2	17753	2	CQ790449	CQ790449 Sequence
C 868	13.8	81.2	9846	5	AL359506	AX08339 HSV-1 (stra	C 941	13.8	81.2	18116	2	CS165132	CS165132 Sequence
C 869	13.8	81.2	9871	2	AR215120	AX08339 HSV-1 (stra	C 942	13.8	81.2	18365	11	AY138246	AY138246 Gallus su
C 870	13.8	81.2	9871	2	AR302362	AX08339 HSV-1 (stra	C 943	13.8	81.2	18495	15	BSU93875	BSU93875 Bacillus su
C 871	13.8	81.2	9871	2	AR373234	AX08339 HSV-1 (stra	C 944	13.8	81.2	20216	2	CQ363808	CQ363808 Sequence
C 872	13.8	81.2	9871	2	AR401620	AX08339 HSV-1 (stra	C 945	13.8	81.2	22183	12	AC014957	AC014957 Drosophila
C 873	13.8	81.2	9871	2	AR442953	AX08339 HSV-1 (stra	C 946	13.8	81.2	22478	15	AB113568	AB113568 Actinopia
C 874	13.8	81.2	9871	2	AR568595	AX08339 HSV-1 (stra	C 947	13.8	81.2	24102	12	AC018225	AC018225 Drosophila
C 875	13.8	81.2	9871	2	AR568595	AX08339 HSV-1 (stra	C 948	13.8	81.2	25760	2	AX708628	AX708628 Sequence
C 876	13.8	81.2	10060	2	AR215121	AX08339 HSV-1 (stra	C 949	13.8	81.2	26550	5	AF421855	AF421855 Homo sapi
C 877	13.8	81.2	10060	2	AR215121	AX08339 HSV-1 (stra	C 950	13.8	81.2	27251	5	AX284661	AX284661 Human DNA
C 878	13.8	81.2	10060	2	AR302363	AX08339 HSV-1 (stra	C 951	13.8	81.2	27435	2	CQ873891	CQ873891 Sequence
C 879	13.8	81.2	10060	2	AR373235	AX08339 HSV-1 (stra	C 952	13.8	81.2	27435	5	AF535142	AF535142 Homo sapi
C 880	13.8	81.2	10060	2	AR401621	AX08339 HSV-1 (stra	C 953	13.8	81.2	27652	5	AF495910	AF495910 Homo sapi
C 881	13.8	81.2	10060	2	AR442954	AX08339 HSV-1 (stra	C 954	13.8	81.2	28058	2	AR619857	AR619857 Sequence
C 882	13.8	81.2	10060	2	AR568596	AX08339 HSV-1 (stra	C 955	13.8	81.2	28762	2	AR619837	AR619837 Sequence
C 883	13.8	81.2	10188	15	AE012274	AX08339 HSV-1 (stra	C 956	13.8	81.2	29183	12	AC014153	AC014153 Drosophila
C 884	13.8	81.2	10222	6	D89801	AX08339 HSV-1 (stra	C 957	13.8	81.2	29793	5	DQ173606	DQ173606 Homo sapi
C 885	13.8	81.2	10285	2	AX551315	AX08339 HSV-1 (stra	C 958	13.8	81.2	29899	2	AR619870	AR619870 Sequence
C 886	13.8	81.2	10285	2	AX552015	AX08339 HSV-1 (stra	C 959	13.8	81.2	30464	5	DQ173597	DQ173597 Homo sapi
C 887	13.8	81.2	10286	15	AE012336	AX08339 HSV-1 (stra	C 960	13.8	81.2	32846	5	AX647129	AX647129 Sequence
C 888	13.8	81.2	10325	2	AR210352	AX08339 HSV-1 (stra	C 961	13.8	81.2	32972	5	DQ173593	DQ173593 Homo sapi
C 889	13.8	81.2	10330	2	CQ789661	AX08339 HSV-1 (stra	C 962	13.8	81.2	32972	5	DQ173598	DQ173598 Homo sapi
C 890	13.8	81.2	10477	2	CQ789659	AX08339 HSV-1 (stra	C 963	13.8	81.2	32972	5	DQ173611	DQ173611 Homo sapi
C 891	13.8	81.2	10516	2	CQ789657	AX08339 HSV-1 (stra	C 964	13.8	81.2	32975	5	DQ173592	DQ173592 Homo sapi
C 892	13.8	81.2	10551	2	AX577781	AX08339 HSV-1 (stra	C 965	13.8	81.2	32975	5	DQ173594	DQ173594 Homo sapi
C 893	13.8	81.2	10561	2	CQ789655	AX08339 HSV-1 (stra	C 966	13.8	81.2	32975	5	DQ173595	DQ173595 Homo sapi
C 894	13.8	81.2	10615	2	CQ789682	AX08339 HSV-1 (stra	C 967	13.8	81.2	32975	5	DQ173596	DQ173596 Homo sapi


```
968 13.8 81.2 32975 5 DQ173599 Homo sapi
969 13.8 81.2 32975 5 DQ173600 Homo sapi
970 13.8 81.2 32975 5 DQ173602 Homo sapi
971 13.8 81.2 32975 5 DQ173603 Homo sapi
972 13.8 81.2 32975 5 DQ173604 Homo sapi
973 13.8 81.2 32975 5 DQ173605 Homo sapi
974 13.8 81.2 32975 5 DQ173607 Homo sapi
975 13.8 81.2 32975 5 DQ173608 Homo sapi
976 13.8 81.2 32975 5 DQ173609 Homo sapi
977 13.8 81.2 32975 5 DQ173610 Homo sapi
978 13.8 81.2 32975 5 DQ173612 Homo sapi
979 13.8 81.2 32975 5 DQ173613 Homo sapi
980 13.8 81.2 32975 5 DQ173614 Homo sapi
981 13.8 81.2 32975 5 DQ173615 Homo sapi
982 13.8 81.2 32975 5 DQ173616 Homo sapi
983 13.8 81.2 32975 5 DQ173618 Homo sapi
984 13.8 81.2 32975 5 DQ173619 Homo sapi
985 13.8 81.2 32975 5 DQ173620 Homo sapi
986 13.8 81.2 32975 5 DQ173621 Homo sapi
987 13.8 81.2 32975 5 DQ173623 Homo sapi
988 13.8 81.2 32975 5 DQ173624 Homo sapi
989 13.8 81.2 32975 5 DQ173625 Homo sapi
990 13.8 81.2 32975 5 DQ173626 Homo sapi
991 13.8 81.2 32975 5 DQ173627 Homo sapi
992 13.8 81.2 32975 5 DQ173628 Homo sapi
993 13.8 81.2 32975 5 DQ173629 Homo sapi
994 13.8 81.2 32975 5 DQ173630 Homo sapi
995 13.8 81.2 32975 5 DQ173631 Homo sapi
996 13.8 81.2 32975 5 DQ173632 Homo sapi
997 13.8 81.2 32975 5 DQ173633 Homo sapi
998 13.8 81.2 32975 5 DQ173634 Homo sapi
999 13.8 81.2 32975 5 DQ173635 Homo sapi
1000 13.8 81.2 32975 5 DQ173636 Homo sapi
```

ALIGNMENTS

```
RESULT 1
BD233463 603 bp DNA linear PAT 17-JUL-2003
LOCUS Human protein having hydrophobic domain and DNA encoding the same.
DEFINITION BD233463
ACCESSION BD233463.1 GI:33043233
VERSION JP 2002519016-A/9.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 603)
AUTHORS Kato,S. and Kimura,T.
TITLE Human protein having hydrophobic domain and DNA encoding the same
JOURNAL Patent: JP 2002519016-A 9 02-JUL-2002;
SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
COMMENT OS Homo sapiens (human)
PN JP 2002519016-A/9
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO,TOMOKO KIMURA
PC C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
FH Key same Location/Qualifiers
FT source 1..603
FT /organism='Homo sapiens (human)'.
FT Location/Qualifiers
1..603
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
FEATURES
source
```

```
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 603;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
Db 550 GCCCAGCGTTGGCCGAG 566
RESULT 2
BD209699 1212 bp DNA linear PAT 17-JUL-2003
LOCUS Compositions isolated from skin cells and methods for their use.
DEFINITION BD209699
ACCESSION BD209699
VERSION BD209699.1 GI:33019469
KEYWORDS JP 2002512798-A/171.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Strachan,L., Sleeman,M., Watson,J.D., Onrust,R., Kumble,A. and
Murison,J.G.
TITLE Compositions isolated from skin cells and methods for their use
JOURNAL Patent: JP 2002512798-A 171 08-MAY-2002;
GENESIS RESEARCH AND DEVELOPMENT CORP LTD
COMMENT OS Homo sapiens (human)
PN JP 2002512798-A/171
PD 08-MAY-2002
PF 29-APR-1999 JP 2000546009
PR 29-APR-1998 US 09/069726,09-NOV-1998 US 09/188930 PI
LORNA STRACHAN,MATTHEW SLEEMAN,JAMES DOUGLAS WATSON,RENE PI
ONRUST.
PI ANAND KUMBLE,JAMES GREG MURISON
PC C12N15/09,A61K38/00,A61P9/00,A61P17/00,A61P29/00,A61P31/18, PC
A61P35/00,
PC C07K14/47,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12N15/ PC
00,
CC A61K37/02,C12N5/00
CC Compositions isolated from skin cells and methods for their
FH Key use. Location/Qualifiers
FT source 1..1212
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1212
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 1212;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
Db 990 GCCCAGCGTTGGCCGAG 1006
RESULT 3
AR341505 1212 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 249 from patent US 6573095.
DEFINITION AR341505
ACCESSION AR341505
VERSION AR341505.1 GI:33733640
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
```

Unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Strachan,L.
TITLE Polynucleotides isolated from skin cells
JOURNAL Patent: US 6573095-A 249 03-JUN-2003;
Genesis Research & Development Corporation Limited; Parnell;
NZX;
FEATURES Location/Qualifiers
source 1..1212
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 1212;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 990 GCCCAGCGTTGGCCGAG 1006
RESULT 4
LOCUS AX078375 1619 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 43 from Patent WO0107612.
ACCESSION AX078375
VERSION AX078375.1 GI:13158044
KEYWORDS
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
TITLE Receptors and associated proteins
JOURNAL Patent: WO 0107612-A 43 01-FEB-2001;
Incyte Genomics, Inc. (US).
FEATURES Location/Qualifiers
source 1..1619
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 2681738CB1"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 1137 GCCCAGCGTTGGCCGAG 1153
RESULT 5
LOCUS BD233473 1718 bp DNA linear PAT 17-JUL-2003
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION BD233473
VERSION BD233473.1 GI:33043243
KEYWORDS JP 2002519016-A/19.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1718)
AUTHORS Kato,S. and Kimura,T.
TITLE Human protein having hydrophobic domain and DNA encoding the same
JOURNAL Patent: JP 2002519016-A 19 02-JUL-2002;

SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
OS Homo sapiens (human)
PN JP 2002519016-A/19
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO,TOMOKO KIMURA
PC
C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
same
FH Key 1..1718 Location/Qualifiers
FT source /organism="Homo sapiens (human)".
FEATURES source
1..1718 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 1718;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 561 GCCCAGCGTTGGCCGAG 577
RESULT 6
LOCUS AK172760 1826 bp mRNA linear PRI 07-MAY-2004
DEFINITION Homo sapiens cDNA FLJ23921 fis, clone COL02043, highly similar to
Homo sapiens retinoic acid induced 3 (RAI3).
ACCESSION AK172760
VERSION AK172760.1 GI:47077732
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1826)
AUTHORS Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2004) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:flcdna@mail.ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES Location/Qualifiers
source 1..1826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL02043"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector pME18SFL3"
ORIGIN

```

Query Match      100.0%; Score 17; DB 5; Length 1826;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
Db 638 GCCACGCTTGCCGAG 654

RESULT 7
CQ723177
LOCUS CQ723177 2290 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 9111 from Patent WO02068579.
ACCESSION CQ723177
VERSION CQ723177.1 GI:42284034
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 9111 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
Location/Qualifiers
1..2290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 17; DB 2; Length 2290;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
Db 1117 GCCACGCTTGCCGAG 1133

RESULT 8
BC003665
LOCUS BC003665 2296 bp mRNA linear PRI 24-NOV-2004
DEFINITION Homo sapiens G protein-coupled receptor, family C, group 5, member
A, mRNA (cDNA clone MGC:923 IMAGE:2988011), complete cds.
ACCESSION BC003665
VERSION BC003665.2 GI:33872669
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2296)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

```

```

Buffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2296)
Director MGC Project.
Direct Submission
Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:13177795.
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 3 Row: 1 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 12056470.
FEATURES
source
Location/Qualifiers
1..2296
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:923 IMAGE:2988011"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH MGC_15"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2296
/gene="GPCR5A"
/note="synonyms: RAIG1, GPRCSA, RAI3"
/db_xref="GeneID:9052"
/db_xref="MIM:604138"
95..1168
/gene="GPCR5A"
/codon_start=1
/product="G protein-coupled receptor, family C, group 5,
member A"
/protein_id="AAH03665.1"
/db_xref="GI:1317796"
/db_xref="GeneID:9052"
/db_xref="MIM:604138"
/translation="MATTPDGCNGLKSKYIRLCDKAEANGIVLETATAGVVTSA
FMFLPLVCKVQDSNRKMLPTQFLFLGVIGFGLTFAFLGDGSGTPRFRFLFG
ILFSCIFSLAHAVSLTKLVGRKPLSLVLGLAVGFSLVQDVIAIEYIVLTNRT
NVNVSLSAPRRNEDFVLLTYVFLMALTFMSSFTFCGSGTGWKRGHAIYLTML
LSIAIWAWITLLMLPDFDRWDOTILSSALAANGVFLLAYVSPEFWLLTKQRNPD
YPVEDACKPOLVKKSGVNRAYSQBEITQFEETGDTLYAPYSTHFLQLQNPPQKE
FSIPRAHWAPSPYKDYEVKKEGS"
ORIGIN
Query Match      100.0%; Score 17; DB 5; Length 2296;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
Db 1112 GCCACGCTTGCCGAG 1128

```

```

RESULT 9
AF506289
LOCUS Homo sapiens orphan G protein-coupling receptor PEIG-1 mRNA, PRI 15-JUL-2002
DEFINITION complete cds.
ACCESSION AF506289.1 GI:21779962
VERSION AF506289
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2297)
AUTHORS Cafarella,E.G., Gonzalez-Guerrico,A.M., Costanzo,R., Pivetta,O.H. and
Santa-Coloma,T.A.
TITLE Induced by 12-O-tetradecanoylphorbol-13-acetate (TPA) in T84 human
colon carcinoma cells
JOURNAL Cell. Mol. Biol. 42 (5), 797-804 (1996)
PUBMED 8832110
REFERENCE 2 (bases 1 to 2297)
AUTHORS Cafarella,E.G., Gonzalez-Guerrico,A.M., Costanzo,R., Pivetta,O.H. and
Santa-Coloma,T.A.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2002) Laboratorio de Biologia Celular y
Molecular, Instituto de Investigaciones Bioquimicas Fundacion
Campomar, Patricia Argentina 435, Buenos Aires 1405, Argentina
FEATURES
source
1. 2297
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12.3"
/sex="male"
/cell_line="T84"
/tissue_type="colorectal carcinoma"
/note="derived from dbEST BE519991"
108. .1181
CDS
/note="belongs to the type 3 G protein-coupling receptor
family; characterized by a 7-transmembrane domain motif;
phorbol ester induced protein-1"
/codon_start=1
/product="orphan G protein-coupling receptor PEIG-1"
/protein_id="AAM77594.1"
/db_xref="GI:21779963"
/translation="MATTVPDGRNGLSKYRLCDKAEANGIVLETATAGVVTVA
FMLTPLIVCKVQSNRRKMLPTQFLFLLGVIGFGLTFAFIIGLDGSGTGRFLFG
ILFSICFSLAHAVSLTKLVGRKRPUSLLVILGLAVGFSLVQDVIAIEYIVLTNRT
NVNVFSELSAPRRNEDFVLLTYVFLMALTFILMSFTFCGSGTGWKRHGAH1YLTML
LSIAIWMVITLLMLPDPDRWDPTILSSALAANGWVFLAYVSPFWLLTKQRNPM
YPVEDAFCKPOLKVGVENRAYSQEITQTGFEETGDTLVAPYSTHFLQNQPPQKE
FSIPRAHWPSPYKDYEVKKEGS"
ORIGIN
Query Match 100.0%; Score 17; DB 5; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGCCGAG 17
| | | | | | | | | | | | | | | | | | | | |
Db 1125 GCCACGCTTGCCGAG 1141
| | | | | | | | | | | | | | | | | | | | |

RESULT 10
AX930411
LOCUS Homo sapiens 2302 bp DNA linear PAT 22-DEC-2003
DEFINITION Sequence 2 from Patent WO03087832.
ACCESSION AX930411
VERSION AX930411.1 GI:40312209

```

```

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Terrett,J.A.
TITLE Diagnosis of carcinoma using raig1 polypeptides
JOURNAL Patent: WO 03087832-A 2 23-OCT-2003;
Oxford GlycoSciences (UK) Limited (GB)
FEATURES
source
1. 2302
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGCCGAG 17
| | | | | | | | | | | | | | | | | | | | |
Db 1117 GCCACGCTTGCCGAG 1133
| | | | | | | | | | | | | | | | | | | | |

RESULT 11
AF095448
LOCUS Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA, PRI 29-DEC-1998
DEFINITION complete cds.
ACCESSION AF095448
VERSION AF095448.1 GI:4063889
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2302)
AUTHORS Cheng,Y. and Lotan,R.
TITLE Molecular cloning and characterization of a novel retinoic
acid-inducible gene that encodes a putative G protein-coupled
receptor
JOURNAL J. Biol. Chem. 273 (52), 35008-35015 (1998)
PUBMED 9857033
REFERENCE 2 (bases 1 to 2302)
AUTHORS Cheng,Y. and Lotan,R.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1998) Tumor Biology, The University of Texas M.
D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX
77030, USA
FEATURES
source
1. 2302
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p12.3-p13"
1. 2302
/gene="RAIG1"
/note="induced by retinoic acid"
100. .1173
/codon_start=1
/product="putative G protein-coupled receptor"
/protein_id="AAC98506.1"
/db_xref="GI:4063890"
/translation="MATTVPDGRNGLSKYRLCDKAEANGIVLETATAGVVTVA
FMLTPLIVCKVQSNRRKMLPTQFLFLLGVIGFGLTFAFIIGLDGSGTGRFLFG
ILFSICFSLAHAVSLTKLVGRKRPUSLLVILGLAVGFSLVQDVIAIEYIVLTNRT
NVNVFSELSAPRRNEDFVLLTYVFLMALTFILMSFTFCGSGTGWKRHGAH1YLTML

```

LSIAIWMATLLMLPDRRRDDTILSSALAANGWVFLLAYVSPFWLTKQRNPM
YPVEDAFCKPQLVKSYGVENRAYSQBEITQGPETGDTLYAPYSTHFLQNPQPOKE
FSIPRAHWSPYKDYEVKKEGS"

ORIGIN

Query Match 100.0%; Score 17; DB 5; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 1117 GCCCAGCGTTGGCCGAG 1133

RESULT 12

BD156680

LOCUS BD156680 2446 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156680

VERSION BD156680.1 GI:27862438

KEYWORDS JP 2002191363-A/11523.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 2446)

Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 11523 09-JUL-2002;

COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11523

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU

PI SAITO,

PI JUNTICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof PH

Location/Qualifiers

FT CDS (254) . (1324) .

Location/Qualifiers

1. 2446

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 17; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 1271 GCCCAGCGTTGGCCGAG 1287

RESULT 13

AX877483

LOCUS AX877483 2446 bp DNA linear PAT 17-DEC-2003

DEFINITION Sequence 12388 from Patent EP1074617.

ACCESSION AX877483

VERSION AX877483.1 GI:40032219

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

*ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

TITLE Primers for synthesizing full-length cDNA and their use

JOURNAL Patent: EP 1074617-A 12388 07-FEB-2001;

Research Association for Biotechnology (JP)

FEATURES

source

1. 2446

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

254. 11327

/note="unnamed protein product"

/codon_start=1

/db_xref="GI:40032220"

/translation="MATTVPDGCENGKSKYVRLCDKAEAGIVLETVATAGVVTVA

FMULTPLVCKVQDSNRKMLPTQFLLGVGLGFTAFPIGLDSTGTRPFLFG

ILFSICFCLLAHVAVSLTKLVGRKPLSLVILGLAVGSLVQDVIAIEYIVLTNR

NVNVFSELSAPRRNEFVLLTYVFLMALTFMSSFTGSGTWKRHGAHLYITML

LSIAIWMATLLMLPDRRRDDTILSSALAANGWVFLLAYVSPFWLTKQRNPM

YPVEDAFCKPQLVKSYGVENRAYSQBEITQGPETGDTLYAPYSTHFLQNPQPOKE

FSIPRAHWSPYKDYEVKKEGS"

ORIGIN

Query Match 100.0%; Score 17; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 1271 GCCCAGCGTTGGCCGAG 1287

RESULT 14

AK001761

LOCUS

AK001761

DEFINITION Homo sapiens cDNA FLJ10899 fis, clone NT2RP5003506.

ACCESSION AK001761

VERSION AK001761.1 GI:7023229

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wakatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Togiya, S., Komai, F., Hara, R.,
Hara, H., Tanase, T. O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,

Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., and Kanehori, K.
NEDO human cDNA sequencing project
Unpublished

3 (bases 1 to 2446)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana Kisarazu, Chiba 292-0812, Japan (E-mail: flj-cdn@nifty.com, tel: 81-438-52-3975, fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
Location/Qualifiers
1. .2446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP5003506"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP5"
/note="Cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
254. .1327
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91890.1"
/db_xref="GI:7023230"
/translation="MATTPDGCNRGLKSKYYRLCDKAEAWGIYLETVATAGVVTVA
FMFLPILVCKVQDSNRKMLPTQFLFLGLVGLTFAPIIGLDGSGTRFFLFG
ILFSICFSLAHAVSLTKLVGRKPLSLVLGLAVGFSLVQDVIAIEYIVLTMMRT
NNVFSLSAPRRNEFVLLTLYVLEFMAITFLMSSFTFCGSGTGWKRHGAIHYLTWL
LSIAIYAWITLLMLPDRDWDITLSSALAANGWFFLLAYSPFWLLTKQRNPD
YPVEDAFKQLVKKSIGVENRAYSQEITGFEETGDTLYAPYSTHFLQNLQPPQKE
FSIPRAHAWSPYKDYEVKKEGS"

ORIGIN
Query Match 100.0%; Score 17; DB 5; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS
1 GCCCACGCTTGGCCGAG 17
|||||
Db 1271 GCCCACGCTTGGCCGAG 1287

FEATURES
source
1. .2446
/organism="Homo sapiens"
/mol_type="taxon:9606"
/db_xref="taxon:9606"

REFERENCE
LOCUS DD210040 2456 bp DNA linear PAT 19-JAN-2006
DEFINITION Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer.
ACCESSION DD210040
VERSION DD210040.1 GI:85654022
KEYWORDS JP 2005518782-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2456)
AUTHORS Zlotnik, A., Mack, D.H., Agiz, N., Gish, K.C., Hebeji, P.A., Wilson, K.E. and Afar, D.
TITLE Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer
JOURNAL Patent: JP 2005518782-A 34 30-JUN-2005;
COMMENT PROTEIN DESIGN LABS INC
OS Homo sapiens
PN JP 2005518782-A/34
PD 30-JUN-2005
PF 17-SEP-2002 JP 2003529912
PR 12-APR-2002 US 60/372246, 08-FEB-2002 US 60/355257, PR 08-FEB-2002 US 60/355145, 13-NOV-2001 US 60/350666, PR 20-SEP-2001 US 60/323887, 17-SEP-2001 US 60/323469 PI albert zlotnik, david h mack, natasha agiz, kurt c gish, peter a pi hebeji, pi keith e wilson, daniel afar
CC
FH Key Location/Qualifiers.
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCCGAG 17
|||||
Db 1271 GCCCACGCTTGGCCGAG 1287

REFERENCE
LOCUS CQ981495 2456 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 350 from Patent EP1498424.
ACCESSION CQ981495
VERSION CQ981495.1 GI:58190785
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Bruemendorf, T., Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE Human nucleic acid sequences from lung tumours
JOURNAL Patent: EP 1498424-A 350 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE); Rosenthal, Andre (DE)
FEATURES
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCCGAG 17
|||||
Db 1271 GCCCACGCTTGGCCGAG 1287

RESULT 16
LOCUS DD210040 2456 bp DNA linear PAT 19-JAN-2006
DEFINITION Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer.
ACCESSION DD210040
VERSION DD210040.1 GI:85654022
KEYWORDS JP 2005518782-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2456)
AUTHORS Zlotnik, A., Mack, D.H., Agiz, N., Gish, K.C., Hebeji, P.A., Wilson, K.E. and Afar, D.
TITLE Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer
JOURNAL Patent: JP 2005518782-A 34 30-JUN-2005;
COMMENT PROTEIN DESIGN LABS INC
OS Homo sapiens
PN JP 2005518782-A/34
PD 30-JUN-2005
PF 17-SEP-2002 JP 2003529912
PR 12-APR-2002 US 60/372246, 08-FEB-2002 US 60/355257, PR 08-FEB-2002 US 60/355145, 13-NOV-2001 US 60/350666, PR 20-SEP-2001 US 60/323887, 17-SEP-2001 US 60/323469 PI albert zlotnik, david h mack, natasha agiz, kurt c gish, peter a pi hebeji, pi keith e wilson, daniel afar
CC
FH Key Location/Qualifiers.
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCCGAG 17
|||||
Db 1271 GCCCACGCTTGGCCGAG 1287

```

RESULT 17
AX549168
LOCUS
DEFINITION Sequence 453 from Patent WO02061087.
ACCESSION AX549168
VERSION AX549168.1 GI:25813894
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
TITLE Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
JOURNAL Patent: WO 02061087-A 453 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 1271 GCCCAGCGTTGGCCGAG 1287

RESULT 18
AK122672
LOCUS
DEFINITION Homo sapiens cDNA FLJ16117 fis, clone ASTRO2003632, highly similar
to Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA.
ACCESSION AK122672
VERSION AK122672.1 GI:34527861
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosioki, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yanazaki, M., Nishimura, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotta, T., Kusano, J., Kanehori, K., Kanehori, K., Tanabe, A.,
Hara, H., Tanase, T. O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikeda, Y., Okamoto, S., Okitani, R., Kawakami, T.,

```

```

Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, M., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
JOURNAL
PUBMED 14702039
REFERENCE
2
AUTHORS
Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M.,
Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 3057)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: flj-cdna@nifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: Reverse Proteomics Research Institute, HRI and
RAB.
FEATURES
source
Location/Qualifiers
1. .3057
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="ASTRO2003632"
/cell_type="normal astrocytes (NHA5732)"
/clone_lib="ASTRO2"
/note="cloning vector: pME18SFL3
primary culture, normal astrocytes"
ORIGIN
Query Match 100.0%; Score 17; DB 5; Length 3057;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 1908 GCCCAGCGTTGGCCGAG 1924

RESULT 19
AX188348
LOCUS
DEFINITION Sequence 4043 from Patent WO0142467.
ACCESSION AX188348
VERSION AX188348.1 GI:15139821
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer

```


JOURNAL Patent: WO 0142467-A 4043 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
source 1..3371
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 3371;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAG 17
|||||
Db 1184 GCCACGCTTGCCGAG 1200
RESULT 20
CQ894732 6730 bp DNA linear PAT 05-NOV-2004
LOCUS Sequence 42 from Patent EP1471075.
DEFINITION CQ894732
ACCESSION CQ894732
VERSION CQ894732.1 GI:55467481
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Rosenthal,A.D., Pilarsky,C., Dahl,E., Specht,T., Bruemendorf,T.,
Lichtner,R., Staub,E., Roepcke,S. and Li,X.I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: EP 1471075-A 42 27-OCT-2004;
Hinzmann, Bernd (DE); Rosenthal, Andre (DE); Pilarsky, Christian
(DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie
(DE)
FEATURES Location/Qualifiers
source 1..6730
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 6730;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACGCTTGCCGAG 17
|||||
Db 1271 GCCACGCTTGCCGAG 1287
RESULT 21
AC007688/c 161577 bp DNA linear PRI 30-AUG-2002
LOCUS AC007688
DEFINITION Homo sapiens 12 BAC RPC111-392P7 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
ACCESSION AC007688
VERSION AC007688.15 GI:5815499
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 161577)
AUTHORS Muzny,D., Aronson,A.D., Bouck,J., Bunac,C., Chen,Z., Ding,Y.,
Dugan,S., Durbin,J., Forcum,J., Garcia,C., Gorrell,J.H.,
Gorrell,L.L., Hernandez,J., Issar,A., Jackson,D., Kneitz,S.,
Kondejewski,N., Lau,S., Leal,B., Lee,E., Lichtarge,O., Liu,W.,
Logan,O., Lu,J., Marondel,I., Martinez,C., Merscher,S., Miller,A.,
Montgomery,K., Oswal,G., Pampell,L.R., Parish,B.J., Perez,L.,
Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Shim,C., Simon,M.,
Vo,Q., Williamson,A., Worley,K.C., Xiang,A.M., Yang,R., Yu,W.,
Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 161577)
Worley,K.C.
Direct Submission
Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 1, 1999 this sequence version replaced gi:5757565.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.
ANNOTATION OF FEATURES:
STGs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.
SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.
QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.
QUALSTAT-REPORT-----
----- Summary Statistics -----
Contig length: 161577
Phrap values in estimate: 160751

Average error rate (BCM-Phrap estimate):
Fraction of Phrap values less than 40 :
Number of consensus changing edits:
Number of N's in consensus :

0.000163681
0.0376047
30
0

Consensus changing edits	
Position	Original+Context
7033	acctcgcgt(c)ccgcccctt
47567	aaaaaataa(n)ggaataaat
51135	aaagaagaa(n)aaagaagaa
75582	aaaaaataa(n)aaagatgct
75585	aaaaaataa(n)ggaatgtcgc
75667	ctaaagcaga(n)taagatctta
75752	tttaaatag(g)gcttgcctt
84017	ggaggggaag(g)aaggaaggag
85227	ttttgtttt(n)tggtttttt
92681	ggagttcaag(n)atgcagtgag
111307	aatctcttat(n)ccgaaattca
111374	actagtatac(n)atcnttttt
111378	gtatacnatc(n)ttttttttt
112621	aaaaaataa(n)ccatcctaga
135812	attcacctc(c)ttttttttt
137207	tgcaggcac(n)cgccaccacg
137218	cgccaccacg(n)ctggctaagt
145113	gcaagtgaan(n)natgtagaat
145114	caagtgaan(n)atgtagaata
145232	ggcaccggtg(n)ntcacnccag
145233	gcaccggtg(n)tcacnccagt
145238	ggcgnntcac(n)ccagtaatcc
145321	gcaacatggt(t)nncccccattc
145322	caacatggtt(n)nncccccattc
145323	acacatggtt(n)nncccccattc
145324	acacatggtt(n)nncccccattc
145363	gtcacagg(n)gtggtgcgt
145377	gaggcgtgta(t)ttgtagcct
145397	tagctacttg(g)gaggaggagt
145505	aaaaaaaaa(n)gaaaaaaaa

----- Distribution of Quality < 40 Bases -----

# bases	Phrap Value Range									
	5	10	15	20	25	30	35	40		
1000	*	*	*	*	*	*	*	*	*	*
900	*	*	*	*	*	*	*	*	*	*
800	*	*	*	*	*	*	*	*	*	*
700	*	*	*	*	*	*	*	*	*	*
600	*	*	*	*	*	*	*	*	*	*
500	*	*	*	*	*	*	*	*	*	*
400	*	*	*	*	*	*	*	*	*	*
300	*	*	*	*	*	*	*	*	*	*
200	*	*	*	*	*	*	*	*	*	*
100	*	*	*	*	*	*	*	*	*	*
0	*	*	*	*	*	*	*	*	*	*

Version: 1.01 gxf.
Location/Qualifiers
1. .161577
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-392P7"
complement(3. .65)
/rpt_family="MER4B"
161. .471
/rpt_family="AluSx"
complement(607. 702)
/rpt_family="MIR"

repeat_region
repeat_region
repeat_region
STS
repeat_region
repeat_region
repeat_region
repeat_region
misc_feature
STS
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

703. .838
/rpt_family="FLAM C"
complement(839. .916)
/rpt_family="MIR"
complement(1205. .1377)
/rpt_family="MER104"
1388. .1525
/standard_name="WIAF-759-STS"
/db_xref="dbSTS:65358"
1683. .1971
/rpt_family="AluJo"
1974. .2261
/rpt_family="AluJb"
2486. .2593
/rpt_family="MIR"
2610. .3214
/note="Region: UniGene cluster similar to AA056332 and
AI074576"
2712. .2823
/standard_name="SHGC-44583"
/db_xref="dbSTS:48451"
3185. .3491
/rpt_family="AluSg"
complement(6397. .6537)
/rpt_family="L2"
6967. .7158
/rpt_family="CCCG)n"
7094. .7240
/rpt_family="CCCG)n"
complement(7770. .8250)
/rpt_family="MER44B"
8831. .9073
/rpt_family="MER102"
complement(9075. .9303)
/rpt_family="L2"
9671. .10135
/rpt_family="LTR33A"

Query Match 100.0%; Score 17; DB 5; Length 161577;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTGGCCGAG 17
Db 94911 GCCCACGCTGGCCGAG 94895

RESULT 22
AC165037/c

LOCUS
DEFINITION AC165037 213601 bp DNA linear HTG 23-JAN-2006
Bos taurus clone CH240-159B22, WORKING DRAFT SEQUENCE, 25 unordered
pieces.
AC165037
AC165037.2 GI:85665020
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Bos taurus (cattle)

ORGANISM
Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS

1 (bases 1 to 213601)
Muzny,D., Adams,C., Agbai II,O., Allen,C., Albrooks,S., Archer,P.,
Arradondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,
Beraducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
Chacko,J., Chahrour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
Anda,C., Deigado,O., Denson,S., Deramo,C., Ding,Y., Din,H.,
Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
DiIuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,


```

gap      63740..63789
/estimated length=50
gap      81081..81130
/estimated length=50
gap      85684..85725
/estimated length=442
gap      108815..108864
/estimated length=50
gap      126980..127029
/estimated length=50
gap      135626..135675
/estimated length=50
gap      140496..140545
/estimated length=50
gap      145819..145868
/estimated length=50
gap      149187..149236
/estimated length=50
gap      156198..156247
/estimated length=50
gap      160594..160643
/estimated length=50
gap      168169..168218
/estimated length=50
gap      175900..175949
/estimated length=50
gap      177820..177919
/estimated length=unknown
gap      200889..200938
/estimated_length=50

Query Match      100.0%; Score 17; DB 12; Length 213601;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 GCCACGCTTGCCGAG 17
        |||||
Db      83186 GCCACGCTTGCCGAG 83170

RESULT 23
LOCUS    AC166087
DEFINITION Bos taurus clone CH240-103F13, *** SEQUENCING IN PROGRESS ***, 35
unordered pieces.
ACCESSION AC166087
VERSION    AC166087.2 GI:85664213
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE     Bos taurus (cattle)
ORGANISM   Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 284106)
Muzny,D., Adams,C., Agbai II,O., Allen,C., Albrooks,S., Archer,P.,
Arredondo,H., Bandaranaike,D., Bangura,D., Beltran,B., Beltran,R.,
Beraducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
Chacko,J., Chahour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
Cheng,M.-T., Chu,J., Ciers,K., Cockrell,R., Coyle,M., Cree,A.,
Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
Gaskin,C., Gensch,S., Ghose,S., Gill,R., Gonzalez,D.,
Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haeblerlen,K.,
Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B.,
Howell,L.T., Hulyk,S., Hume,J., Ito,K., Jackson,A., Jackson,L.,
Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,

```

```

Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
Li,Y., Li,Z., Linell,M., Liu,Y.-S., Liu,Y., Liyangde,D.,
London,P., Lopez,J., Lorensuewa,L., Lozado,R., Luk,T., Madu,R.,
Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
McLelland,H., McPherson,J., Mercadao,C., Metzker,M.,
Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M.,
Murray,D., Nazarith,L., Ngo,D., Nguyen,N., Norwig-Estaugh,E.,
Nott,A., Nwaokemele,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,
Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quroz,J.,
Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
Trejos,Z., Usmani,K., Vargo,C., Verduzco,D., Villasana,D., Virk,D.,
Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 284106)
Worley,K.C.

Direct Submission
Submitted (23-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 284106)

Bovine Genome Sequencing Consortium
Direct Submission
Submitted (23-JAN-2006) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jan 23, 2006 this sequence version replaced gi:71067167.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: FKHC
Center clone name: CH240-103F13
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 264399 bases at least Q40
Consensus quality: 267053 bases at least Q30
Consensus quality: 269711 bases at least Q20
Estimated insert size: 271703; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one 'clone'.
* NOTE: This is a 'working draft' sequence. It currently
consists of 35 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
3307: contig of 3307 bp in length
3357: gap of 50 bp
3358: 21882: contig of 18525 bp in length
21883: 21932: gap of 50 bp
21933: 27057: contig of 5125 bp in length
27058: 27107: gap of 50 bp
27108: 30386: contig of 3279 bp in length
30387: 30436: gap of 50 bp
30437: 34785: contig of 4349 bp in length
34786: 34912: gap of 127 bp
34913: 37140: contig of 2228 bp in length
37141: 37190: gap of 50 bp
37191: 50210: contig of 13020 bp in length
50211: 50260: gap of 50 bp
60819: 60819: contig of 10559 bp in length
60820: 61482: gap of 663 bp
61483: 64384: contig of 2902 bp in length
64385: 64434: gap of 50 bp
64435: 67014: contig of 2580 bp in length
67015: 67064: gap of 50 bp
67065: 93556: contig of 26492 bp in length
93557: 93606: gap of 50 bp
93607: 108482: contig of 14876 bp in length
108483: 108532: gap of 50 bp
108533: 121591: contig of 13059 bp in length
121592: 121641: gap of 50 bp
121642: 126700: contig of 5059 bp in length
126701: 126750: gap of 50 bp
126751: 128315: contig of 1565 bp in length
128316: 128365: gap of 50 bp
128366: 153628: contig of 25213 bp in length
153629: 153628: gap of 50 bp
153629: 171554: contig of 17926 bp in length
171555: 171604: gap of 50 bp
171605: 179129: contig of 7525 bp in length
179130: 179179: gap of 50 bp
179180: 182152: contig of 2973 bp in length
182153: 18519: gap of 3367 bp
18520: 188536: contig of 3017 bp in length
188537: 188974: gap of 438 bp
188975: 192508: contig of 3534 bp in length
192509: 192559: gap of 50 bp
192560: 197962: contig of 5404 bp in length
197963: 198012: gap of 50 bp
198013: 202007: contig of 3995 bp in length
202008: 202568: gap of 561 bp
202569: 211162: contig of 8594 bp in length
211163: 211212: gap of 50 bp
211213: 247082: contig of 35870 bp in length
247083: 249331: gap of 2249 bp
249332: 253480: contig of 4149 bp in length
253481: 253634: gap of 154 bp
253635: 268332: contig of 14698 bp in length
268333: 268521: gap of 189 bp
268522: 269932: contig of 1411 bp in length
269933: 270032: gap of unknown length
270033: 271073: contig of 1041 bp in length
271074: 271173: gap of unknown length
271174: 272194: contig of 1021 bp in length
272195: 272294: gap of unknown length
272295: 273957: contig of 1663 bp in length
273958: 274057: gap of unknown length
274058: 275360: contig of 1303 bp in length
275361: 275460: gap of unknown length
275461: 276754: contig of 1294 bp in length
276755: 276854: gap of unknown length
276855: 278168: contig of 1314 bp in length
278169: 278268: gap of unknown length
278269: 284106: contig of 5838 bp in length.

FEATURES

source
Location/Qualifiers
1. 284106
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-103F13"
3308. .3357
/estimated_length=50
21883. .21932
/estimated_length=50
27058. .27107
/estimated_length=50
30387. .30436
/estimated_length=50
34786. .34912
/estimated_length=127
37141. .37190
/estimated_length=50
50211. .50260
/estimated_length=50
60820. .61482
/estimated_length=663
64385. .64434
Query Match 100.0%; Score 17; DB 12; Length 284106;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCTTGCCGAG 17
|||||
Db 255390 GCCCAGCTTGCCGAG 255406
RESULT 24
SC0939104/c
LOCUS
DEFINITION Streptomyces coelicolor A3(2) complete genome; segment 1/29.
ACCESSION AL939104 AL109747 AL109949 AL109950 AL109962 AL109972 AL109973
AL109987 AL109989 AL110470 AL117322 AL117385 AL121854 AL590984
AL645882 AL732393
VERSION AL939104.1 GI:24413714
KEYWORDS
SOURCE Streptomyces coelicolor A3(2)
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomycetes.
REFERENCE
1 Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neill, S., Rabinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrall, B.G., Parkhill, J. and Hopwood, D.A.
Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)
Nature 417 (6885), 141-147 (2002)
12000953
2 (bases 1 to 299050)
Bentley, S.D.
Direct Submission
Submitted (09-MAY-2002) Submitted on behalf of the Streptomyces sequencing team, Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: sdb@sanger.ac.uk
On or before Oct 26, 2002 this sequence version replaced
gi:20520931, gi:20520890, gi:20520923, gi:20520881, gi:20520932, gi:20520933, gi:20520892, gi:20520934, gi:20520882, gi:20520771, gi:20520935, gi:13872787, gi:20520986.
Location/Qualifiers
1. .299050
/organism="Streptomyces coelicolor A3(2)"

```

/mol_type="genomic DNA"
/strain="A3 (2)"
/db_xref="taxon:100226"
1. .21653
misc_feature
repeat.
435. .440
446. .1123
/gene="SCO0001"
/note="synonym: SCEND.02c"
446. .1123
/gene="SCO0001"
/note="SCEND.02c, unknown, doubtful CDS, len: 225aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD30875.1"
/db_xref="GI:20520987"
/db_xref="UniProtKB/TrEMBL:Q9FBQ0"
/translacion="MTGHSTGPGTALSSDSTCRVTQYTQAGVNARLRLFALLERRA
CPARVTPWGRKSWTAWRLILGVCCVGRIGRRRGGDGERPGGHRGPGGLATA
RRRSGATELA VHCADVQREADILVRLEGFVRESVLPRAHPHTTARRVLEVLGAG
SLCTARTNSDEDIYLCILGVGHYDDQDPFPKDGKPGWQRAGASINWNGSACIPIHA
AIEGPRK"
1238. .1243
1252. .3813
/gene="SCO0002"
/note="synonyms: SC8E7.42c, SCEND.01c, SCJ24.01c"
1252. .3813
/gene="SCO0002"
/note="SC8E7.42c, SCEND.01c, SCJ24.01c"
1252. .3813
/gene="SCO0002"
/note="SCEND.01c, unknown, partial CDS, len: >27aa;
strongly similar to TR:AAF36550 (EMBL:AF194023)
helicase-like protein from Streptomyces lividans (881 aa)
fasta scores: opt: 114, z-score: 206.0, E(): 0.00052,
80.0% identity in 25 aa overlap.
SCJ24.01c, unknown, partial CDS, len: >457aa; strongly
similar to TR:AAF36550 (EMBL:AF194023) helicase-like
protein from Streptomyces lividans (881 aa) fasta scores;
opt: 2191, z-score: 2490.1, E(): 0, 72.5% identity in 458
aa overlap.
SC8E7.42c, partial CDS, unknown, len:> 431 aa. Highly
similar to Streptomyces lividans
TR:AAF36550 (EMBL:AF194023) helicase-like protein found in
the Chromosome terminal region (881 aa), fasta scores opt:
1940 z-score: 2170.3 E(): 0.66.9% identity in 438 aa
overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD30876.1"
/db_xref="GI:20520988"
/db_xref="GOA:Q8R6K2"
/db_xref="InterPro:IPR001650"
/db_xref="UniProtKB/TrEMBL:Q8R6K2"
/translacion="MGVSTGSGKTTTAAACALEYFPEGRILVMVPTDLIVQSAQSW
RVGHRAPMVAVCSVDKDLVLSQGVRTTNPITQLAWAGTGPVVVFATVAVLVDRE
PADVTGRRVTPGLSALAGGERLGVQRMAPDPAIDLDSGHTMAGDGPWAAIHDS
RIPVDFRLYLTPTKILAPRQGRDRELVIASMEDDSSTYGTIFDLGLAEVER
SLAGFEDLVLEIRDPDITLUSEDALGRNMLAQALLEHAAQNLHMTVWFHORV
EEAFAEQMPRTAAELYTAESAALAEAAELPASSIDAELEYELGHRVPPDRYVA
DWLSCDHP TAHRVITQREANGIDQNRVHRAFLSSVRALGVGVDIAGLRGVEAVCI
VGRSODVNVONI GRALPNPDGTTKTARI11PVLPQEDPKDMVASASYOLPVDI
LQALRSRSMVDQLASRALTRGTERRI.HVPAPAPGPDGMPETSTQEQVEDRV
TSVNVNFA SPDAADIAALTRCVIRPQSLVWLEGYQALVRMRENGITGLYAVPDI
FTEVETKGYPGVQQRHRAHAGELDPHRKELDDAGVMWEPDGEAWENKLAAPF
SYRRAGHLAPQDQAWGEGEAMVPTGOHLNLRKKAAGNLGKDPERAAERAAQLAA
IPDNPCWPLDQHRVYRLADLVADGVLPADIPGVLPFGDDIGKDPTRORASTWT
OLSAEQBELTALGVTPVERPAPAPASTSPKSPKAKOSFORGLAALQWVERGAD
RVPFRHASEIADVEADPVIILGLWVSNTRARRDKLTAEQLNALRGLGWEMAA"
3869. .6220
/gene="SCO0003"
/note="synonym: SC8E7.41c"

```

SLGRPELTIAVDATRSILAAVLPRYSTKAVDAALLAEMAVHPHARTWPSALHSR
REQVPERLSLDERLEGAARPVVVPETIVDRGKIYLSQGFVAACMLGVSVQAPP
ASPOKAVVERTFEGAINDLFCOHVGHGTSNPQRGLTTAETRTWITPOLQDFLEWI
TCQWNRHDLGRHVPFKTALTPNQMAALITSGYFVPLSGADYLELLPVRWQPI
TERGLRDYRTYNNHDLPHRGQSFVASKDGKWEVHNPHDARQIFVRLTDGQJHEI
PWIRHDVHQPFNEAIWRHVQAEVQRGRDQHEADLADALDOLLRRTHLAETQKT
RRRATRSQTAAQLPDLPCQRRPFAETAPAPAPDWSLSDDLISVDTAAQTGTSEME
GASVPPAEAGGYGLWDAEAEQW"
8231..8623
/gene="SC00005"
/note="Prim match to entry PF00665 rve, Integrase core
domain, score 47.20, E-value 1.8e-12"
9304..10410
/gene="SC00006"
/note="synonyms: SC1C9.03, SCJ30.01"
9304..10410

misc_feature

gene

CDS

Query Match 100.0%; Score 17; DB 15; Length 299050;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAG 17
|||||

Db 216438 GCCACGCTTGCCGAG 216422

RESULT 25

CP000124_29

WPCOMMENT

Sequence split into 42 fragments LOCUS CP000124 Accession CP000124

Fragment Name	Begin	End
CP000124_00	1	110000
CP000124_01	100001	210000
CP000124_02	200001	310000
CP000124_03	300001	410000
CP000124_04	400001	510000
CP000124_05	500001	610000
CP000124_06	600001	710000
CP000124_07	700001	810000
CP000124_08	800001	910000
CP000124_09	900001	1010000
CP000124_10	1000001	1110000
CP000124_11	1100001	1210000
CP000124_12	1200001	1310000
CP000124_13	1300001	1410000
CP000124_14	1400001	1510000
CP000124_15	1500001	1610000
CP000124_16	1600001	1710000
CP000124_17	1700001	1810000
CP000124_18	1800001	1910000
CP000124_19	1900001	2010000
CP000124_20	2000001	2110000
CP000124_21	2100001	2210000
CP000124_22	2200001	2310000
CP000124_23	2300001	2410000
CP000124_24	2400001	2510000
CP000124_25	2500001	2610000
CP000124_26	2600001	2710000
CP000124_27	2700001	2810000
CP000124_28	2800001	2910000
CP000124_29	2900001	3010000
CP000124_30	3000001	3110000
CP000124_31	3100001	3210000
CP000124_32	3200001	3310000
CP000124_33	3300001	3410000
CP000124_34	3400001	3510000
CP000124_35	3500001	3610000
CP000124_36	3600001	3710000
CP000124_37	3700001	3810000
CP000124_38	3800001	3910000
CP000124_39	3900001	4010000
CP000124_40	4000001	4110000
CP000124_41	4100001	426292

Continuation (30 of 42) of CP000124 from base 2900001 (CP000124 Burkholderia pseudomallei:

Query Match 94.1%; Score 16; DB 15; Length 110000;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
|||||

Db 43511 CCCACGCTTGCCGAG 43526

RESULT 26

BX571965_26

WPCOMMENT

Sequence split into 41 fragments LOCUS BX571965 Accession BX571965

Fragment Name	Begin	End
BX571965_00	1	110000
BX571965_01	100001	210000
BX571965_02	200001	310000
BX571965_03	300001	410000
BX571965_04	400001	510000
BX571965_05	500001	610000
BX571965_06	600001	710000
BX571965_07	700001	810000
BX571965_08	800001	910000
BX571965_09	900001	1010000
BX571965_10	1000001	1110000
BX571965_11	1100001	1210000
BX571965_12	1200001	1310000
BX571965_13	1300001	1410000
BX571965_14	1400001	1510000
BX571965_15	1500001	1610000
BX571965_16	1600001	1710000
BX571965_17	1700001	1810000
BX571965_18	1800001	1910000
BX571965_19	1900001	2010000
BX571965_20	2000001	2110000
BX571965_21	2100001	2210000
BX571965_22	2200001	2310000
BX571965_23	2300001	2410000
BX571965_24	2400001	2510000
BX571965_25	2500001	2610000
BX571965_26	2600001	2710000
BX571965_27	2700001	2810000
BX571965_28	2800001	2910000
BX571965_29	2900001	3010000
BX571965_30	3000001	3110000
BX571965_31	3100001	3210000
BX571965_32	3200001	3310000
BX571965_33	3300001	3410000
BX571965_34	3400001	3510000
BX571965_35	3500001	3610000
BX571965_36	3600001	3710000
BX571965_37	3700001	3810000
BX571965_38	3800001	3910000
BX571965_39	3900001	4010000
BX571965_40	4000001	4074542

Continuation (27 of 41) of BX571965 from base 2600001 (BX571965 Burkholderia pseudomallei:

Query Match 94.1%; Score 16; DB 15; Length 110000;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
|||||

Db 73449 CCCACGCTTGCCGAG 73464

RESULT 27

CP000010_16

WPCOMMENT

Sequence split into 36 fragments LOCUS CP000010 Accession CP000010

Fragment Name	Begin	End
---------------	-------	-----


```

CP000010_00      1 110000
CP000010_01      100001
CP000010_02      210000
CP000010_03      300001
CP000010_04      400001
CP000010_05      500001
CP000010_06      600001
CP000010_07      700001
CP000010_08      800001
CP000010_09      900001
CP000010_10     1000001
CP000010_11     1100000
CP000010_12     1200000
CP000010_13     1300000
CP000010_14     1400001
CP000010_15     1500001
CP000010_16     1600001
CP000010_17     1700001
CP000010_18     1800001
CP000010_19     1900001
CP000010_20     2000001
CP000010_21     2100000
CP000010_22     2200001
CP000010_23     2300001
CP000010_24     2400001
CP000010_25     2500001
CP000010_26     2600001
CP000010_27     2700001
CP000010_28     2800001
CP000010_29     2900001
CP000010_30     3000001
CP000010_31     3100001
CP000010_32     3200001
CP000010_33     3300001
CP000010_34     3400001
CP000010_35     3500001
CP000010_36     350148
Continuation (17 of 36) of CP000010 from base 1600001 (CP000010 Burkholderia mallei ATCC

```

Query Match 94.1%; Score 16; DB 15; Length 110000;

Best Local Similarity 100.0%; Pred. No. 9.5e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17

Db 98066 CCCACGCTTGCCCGAG 98081

RESULT 28

AC167299/c

LOCUS AC167299 174298 bp DNA linear HTG 22-OCT-2005
DEFINITION Oryctolagus cuniculus clone LBL-35017, WORKING DRAFT SEQUENCE, 6
ordered pieces.

AC167299

VERSION AC167299.3 GI:78042356

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE Oryctolagus cuniculus (rabbit)

ORGANISM Oryctolagus cuniculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha;

Leporidae; Oryctolagus.

1 (bases 1 to 174298).

Antonellis,A., Ayele,K., Bass,D., Benjamin,B., Bera,J.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G.,
Coleman,H., Franks,S., Fuksenko,T., Gestole,M., Greene,A., Guan,X.,
Gupta,J., Gurson,N., Haghighi,P., Han,E., Han,J., Hansen,N.,
Ho,S.-I., Hu,P., Hunter,G., Hurle,B., Idol,J.R., Kwong,P.,
Laric,P., Larson,S., Lee-lin,S.-Q., Legaspi,R., Madden,M.,
Maduro,Q.D., Maduro,V.B., Margulies,E.H., Mastello,C., Maskery,B.,
McDowell,J., Montemayor,C., Mullikin,J.C., Park,M., Portnoy,M.E.,
Prasad,A., Puri,O., Rantz,K., Reddix-Dugue,N., Sante,A.,
Schandler,K., Schueler,M.G., Sison,C., Stantropop,S., Tave,A.,
Thomas,J.W., Thomas,P.J., Tsipouri,V., Ung,L., Vogt,J.L.,
Wetherby,K.D., Withers,T.R., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 174298)

Green,E.D.

Direct Submission

Submitted (24-AUG-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA

3 (bases 1 to 174298)

Green,E.D.

Direct Submission

Submitted (22-OCT-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA

On Oct 22, 2005 this sequence version replaced gi:77627864.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: lnj

Center clone name: 035017

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is generally based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 172944 bases at least Q40

Consensus quality: 173403 bases at least Q30

Consensus quality: 173643 bases at least Q20

Insert size: 176000; agarose-fp

Insert size: 173798; sum-of-contigs

Quality coverage: 9.18x in Q20 bases; agarose-fp

Quality coverage: 9.29x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and * the accession number will be preserved.

* 1 56412: contig of 56412 bp in length

* 56413 56512: gap of unknown length

* 56513 61819: contig of 5307 bp in length

* 61820 61919: gap of unknown length

* 61920 68051: contig of 6132 bp in length

* 68052 68151: gap of unknown length

* 68152 101952: contig of 33801 bp in length

* 101953 102052: gap of unknown length

* 102053 157436: contig of 55384 bp in length

* 157437 157536: gap of unknown length

* 157537 174298: contig of 16762 bp in length.

Location/Qualifiers

1..174298

/organism="Oryctolagus cuniculus"

/mol_type="genomic DNA"

/db_xref="taxon:9986"

/clone="LBL-35017"

/clone_lib="LB1"

FEATURES

source

/note="BAC resource: http://bacpac.chori.org/
breed: New Zealand White"

misc_feature

1..56412
/note="assembly_fragment
clone_end:17
vector_side:left"

gap

56413..56512
/estimated_length=unknown

misc_feature

56513..61819
/note="assembly_fragment"

gap

61820..61919
/estimated_length=unknown

misc_feature

61920..68051
/note="assembly_fragment"

gap

68052..68151
/estimated_length=unknown

misc_feature

68152..101952
/note="assembly_fragment"

gap

101953..102052
/estimated_length=unknown

misc_feature

102053..157436
/note="assembly_fragment"

gap

157437..157536
/estimated_length=unknown

misc_feature

157537..174298
/note="assembly_fragment
vector_side:right"

ORIGIN

Query Match 94.1%; Score 16; DB 12; Length 174298;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16

Db 87377 GCCCAGCGTTGGCCGA 87362

RESULT 29

MAU564400/c MAU564400 117 bp DNA linear BCT 15-APR-2005
LOCUS Mycobacterium aurum partial GyrB gene for DNA gyrase subunit B,
DEFINITION strain ATCC 23366.

ACCESSION AJ564400.1 GI:31559458

VERSION DNA gyrase subunit B; GyrB gene.

KEYWORDS Mycobacterium aurum

SOURCE Mycobacterium aurum

ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.

REFERENCE

1 Dauendorffer, J.N., Guillemin, I., Aubry, A., Truffot-Pernot, C.,
Sougakoff, W., Jarlier, V. and Cambau, E.

TITLE Identification of mycobacterial species by PCR sequencing of
quinolone resistance-determining regions of DNA gyrase genes

J. Clin. Microbiol. 41 (3), 1311-1315 (2003)

PUBMED 12624075

REFERENCE 2 (bases 1 to 117)

AUTHORS Cambau, E.

TITLE Direct Submission

JOURNAL Submitted (26-FEB-2003) Cambau E., Bacteriologie, Faculte de
Medecine Pitie-Salpetriere, 91 Boulevard de l'Hopital, 75013 Paris,
FRANCE

FEATURES

source Location/Qualifiers

1..117
/organism="Mycobacterium aurum"
/mol_type="genomic DNA"
/strain="ATCC 23366"
/db_xref="taxon:1791"
<1..>117
/gene="GyrB"
<1..>117

gene

CDS

/gene="GyrB"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="CAD92772.1"
/db_xref="GI:31559458"
/db_xref="UniProtKB/TREMBL:Q7WTA2"
/translation="DSAGGSAKSGRDSMFQAILPLRGKIINVEKARIDRVLKN"

ORIGIN

Query Match 90.6%; Score 15.4; DB 15; Length 117;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 31 GCCCAGCTTGGCCGAG 15

RESULT 30

MAU181332 MAU181332 385 bp DNA linear PRI 11-JAN-2006
LOCUS Homo sapiens partial HLA-D gene for MHC class II antigen,
DEFINITION HLA-DQB1*06new allele, exon 2.
ACCESSION AM181332

VERSION AM181332.1 GI:84794315

KEYWORDS HLA-D gene; HLA-DQB1*06new allele; human leucocyte antigen D; major
histocompatibility complex; MHC class II antigen.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE

AUTHORS Witter, K.

TITLE A novel HLA-DQB1 allele, DQB1*06xx was detected through routine HLA
class II high resolution typing

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 385)

AUTHORS Witter, K.

TITLE Direct Submission

JOURNAL Submitted (10-JAN-2006) Witter K., Labor of Immunogenetics,
University Munich, Paul-Heyse Strasse 33, 80336 Munich, GERMANY

FEATURES

source Location/Qualifiers

1..385
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="158308"
/db_xref="taxon:9606"
<1..>283
/gene="HLA-D"
<1..>283
/allele="HLA-D"
/allele="HLA-DQB1*06new"
<1..>283
/gene="HLA-D"
/function="antigen presentation"
/function="antigen presenting molecule"
/codon_start=3
/allele="HLA-DQB1*06new"
/product="MHC class II antigen"
/protein_id="CAJ57391.1"
/db_xref="GI:84794316"
/translation="DFVQFKGCMYFTNGTERVLVTRHIYNREYVRFDSVDGVYRA
VTPQGRPAEYWNYSQKVLGTRAELEDTVCHHNYEVAFRGILQRRGRRR"
1..283
/gene="HLA-D"
/number=2
/allele="HLA-DQB1*06new"

ORIGIN

Query Match 90.6%; Score 15.4; DB 5; Length 385;
Best Local Similarity 94.1%; Pred. No. 1.1e+03;

```

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
    ||||| ||||| |||||
Db 296 GCCCAGCGTTGGCCGAG 312

RESULT 31
LOCUS AY535046/c 586 bp DNA linear ENV 01-MAR-2004
DEFINITION Uncultured eukaryote clone E8 18S ribosomal RNA gene, partial
sequence.
ACCESSION AY535046
VERSION AY535046.1 GI:42794960
KEYWORDS ENV.
SOURCE uncultured eukaryote
ORGANISM uncultured eukaryote
REFERENCE 1 (bases 1 to 586)
AUTHORS Duplessis,M.R., Dufour,S.C., Blankenship,L.E., Yayanos,A.A. and Felbeck,H.
CONSTRM SIO
TITLE Anatomical and Experimental Evidence for Particulate Feeding in Lucinoma aequizonata and Parvilucina tenuisculpta (Bivalvia: Lucinidae) from the Santa Barbara Basin
JOURNAL Marine Biology (2004) In press
REFERENCE 2 (bases 1 to 586)
AUTHORS Blankenship,L.E., Duplessis,M.R., Dufour,S.C., Yayanos,A.Aristides. and Felbeck,H.
CONSTRM SIO
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2004) Marine Biology Research Division, Scripps Institution of Oceanography, 9500 Gilman Dr. 0202, La Jolla, CA 92093, USA
FEATURES
source 1..586
    /organism="uncultured eukaryote"
    /mol_type="genomic DNA"
    /isolation_source="Lucinoma aequizonata gut"
    /db_xref="taxon:100272"
    /clones="E8"
    /environmental sample
    complement(<1..586)
    /product="18S ribosomal RNA"

rRNA

ORIGIN
Query Match 90.6%; Score 15.4; DB 1; Length 586;
Best Local Similarity 94.1%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
    ||||| ||||| |||||
Db 283 GCCCAGCGTTGGCCGAG 267

RESULT 32
LOCUS DQ288979/c 738 bp DNA linear VRL 05-DEC-2005
DEFINITION Bovine leukemia virus strain 991 integrase-like (pol) gene, partial
sequence.
ACCESSION DQ288979
VERSION DQ288979.1 GI:82791298
KEYWORDS Bovine leukemia virus
SOURCE Bovine leukemia virus
ORGANISM Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae; Deltaretrovirus.
REFERENCE 1 (bases 1 to 738)
AUTHORS Hirsch,C., Barbosa-Stancioli,E.F., Camargos,M.F., Reis,J.K.P. and Leite,R.C.
TITLE Bovine leukemia virus pol gene: Genetic variability and phylogeny of the integrase region of Brazilian samples
JOURNAL Unpublished

```

```

REFERENCE 2 (bases 1 to 738)
AUTHORS Hirsch,C., Barbosa-Stancioli,E.F., Camargos,M.F., Reis,J.K.P. and Leite,R.C.
TITLE Direct Submission
JOURNAL Veterinaria da Universidade Federal de Minas Gerais, Escola de Veterinaria da Universidade Federal de Minas Gerais, Avenida Antonio Carlos 6627. Bairro Sao Francisco caixa postal 567., Belo Horizonte, Minas Gerais 30.123-970, Brasil
FEATURES
source 1..738
    /organism="Bovine leukemia virus"
    /proviral
    /mol_type="genomic DNA"
    /strain="991"
    /specific_host="Bos taurus"
    /db_xref="taxon:11901"
    /country="Brazil; Pedro Leopoldo, Minas Gerais State"
    <1..738
    /gene="pol"
    /misc_feature <1..738
    /genes="pol"
    /note="similar to integrase"

ORIGIN
Query Match 90.6%; Score 15.4; DB 10; Length 738;
Best Local Similarity 94.1%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
    ||||| ||||| |||||
Db 188 GCCCAGCGTTGGCCGAG 172

RESULT 33
LOCUS AY381191 1802 bp DNA linear ENV 08-JUN-2004
DEFINITION Uncultured eukaryote clone BL000921.11 18S rRNA gene, complete
sequence.
ACCESSION AY381191
VERSION AY381191.1 GI:39981844
KEYWORDS ENV.
SOURCE uncultured marine eukaryote
ORGANISM uncultured marine eukaryote
REFERENCE 1 (bases 1 to 1802)
AUTHORS Massana,R., Castresana,J., Balague,V., Guillou,L., Romari,K., Groisillier,A., Valentin,K. and Pedros-Allo,C.
TITLE Phylogenetic and Ecological Analysis of Novel Marine Stramenopiles Appl. Environ. Microbiol. 70 (6), 3528-3534 (2004)
JOURNAL PUBMED 15184153
REFERENCE 2 (bases 1 to 1802)
AUTHORS Massana,R., Castresana,J., Balague,V., Guillou,L., Romari,K., Groisillier,A., Valentin,K. and Pedros-Allo,C.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2003) Marine Biology and Oceanography, Institut de Ciències del Mar, CSIC, CSIC, Passeig Marítim de la Barceloneta 37-49, Barcelona, Catalonia 08003, Spain
FEATURES
source 1..1802
    /organism="uncultured marine eukaryote"
    /mol_type="genomic DNA"
    /isolation_source="coastal surface water, Northwestern Mediterranean, Banes Bay"
    /db_xref="taxon:203449"
    /clones="BL000921.11"
    /environmental sample
    /product="18S ribosomal RNA"

rRNA

ORIGIN
Query Match 90.6%; Score 15.4; DB 1; Length 1802;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY 1 GCCCAGCTTGGCCGAG 17
Db 1496 GCACACGCTTGGCCGAG 1512

RESULT 34
AX416413/c
LOCUS AX416413 2306 bp DNA linear PAT 02-SEP-2002
DEFINITION Sequence 3404 from Patent WO0228891.
ACCESSION AX416413
VERSION AX416413.1 GI:21448870
KEYWORDS
SOURCE Listeria monocytogenes ATCC 19115
ORGANISM Listeria monocytogenes ATCC 19115
REFERENCE 1
AUTHORS Kunst, F. and Glaser, P.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 3404 11-APR-2002;
INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES
source
1..2306
/organism="Listeria monocytogenes ATCC 19115"
/mol_type="unassigned DNA"
/db_xref="taxon:176281"

ORIGIN
Query Match 90.6%; Score 15.4; DB 2; Length 2306;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
Db 834 GCCCAGCTTGGCTGAG 818

RESULT 35
HSA564991
LOCUS HSA564991 2550 bp DNA linear PRI 27-JUL-2005
DEFINITION Homo sapiens HLA-DQB1 gene for MHC class II antigen, HLA-DQB1*0602
allele, intron 2.
ACCESSION AJ564991
VERSION AJ564991.1 GI:31455385
KEYWORDS HLA-DQB1 gene; HLA-DQB1*0602 allele; major histocompatibility
complex; MHC class II antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Dunn, P.P., Day, S., Williams, S. and Bendukidze, N.
TITLE HLA-DQB1 sequencing-based typing using newly identified conserved
nucleotide sequences in introns 1 and 2
JOURNAL Tissue Antigens 66 (2), 99-106 (2005)
PUBMED 16029429
AUTHORS Dunn, P.P.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2003) Dunn P.P., H & I DNA Reference Laboratory,
National Blood Service, Southmead Road, BS10 5ND, UNITED KINGDOM
FEATURES
source
1..2550
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_line="W6V"
1..2550
/gene="HLA-DQB1"
1..2550
/gene="HLA-DQB1"
/number=2
/functions="antigen presenting molecule"
/allele="HLA-DQB1*0603"

ORIGIN
Query Match 90.6%; Score 15.4; DB 5; Length 2550;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
Db 26 GCCCAGCTTGGCCGAG 42

RESULT 37
HSA564993
LOCUS HSA564993 2550 bp DNA linear PRI 27-JUL-2005
DEFINITION Homo sapiens HLA-DQB1 gene for MHC class II antigen,
HLA-DQB1*060401 allele, intron 2.
ACCESSION AJ564993
VERSION AJ564993.1 GI:31455387
KEYWORDS HLA-DQB1 gene; HLA-DQB1*060401 allele; major histocompatibility
complex; MHC class II antigen.
SOURCE Homo sapiens (human)
```

ORGANISM	Homo sapiens	JOURNAL	Proc Natl Acad Sci U S A 99 (26), 16899-16903 (2002)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	PUBMED	12477932
AUTHORS	1 Dunn, P.P., Day, S., Williams, S. and Bendukidze, N.	REFERENCE	2 (bases 1 to 2623)
TITLE	HUA-DQB1 sequencing-based typing using newly identified conserved nucleotide sequences in introns 1 and 2	AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
JOURNAL	Tissue Antigens 66 (2), 99-106 (2005)	TITLE	Direct Submission
PUBMED	16029429	JOURNAL	Submitted (03-JUN-2003) Dunn P.P., H & I DNA Reference Laboratory, National Blood Service, Southmead Road, BS10 5ND, UNITED KINGDOM
REFERENCE	2 (bases 1 to 2550)	REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
AUTHORS	Dunn, P.P.	COMMENT	On Dec 9, 2003 this sequence version replaced gi:17511964.
TITLE	Direct Submission	Contact:	MGC help desk
JOURNAL	Submitted (03-JUN-2003) Dunn P.P., H & I DNA Reference Laboratory, National Blood Service, Southmead Road, BS10 5ND, UNITED KINGDOM	Email:	cgaps@remail.nih.gov
FEATURES	Location/Qualifiers	Tissue Procurement:	ATCC
source	1..2550	cDNA Library Preparation:	Rubin Laboratory
	/organism="Homo sapiens"	DNA Sequencing by:	The I.M.A.G.E. Consortium (LLNL)
gene	/mol_type="genomic DNA"	Sequencing Center (NISC),	
intron	/db_xref="taxon:9606"	Gaithersburg, Maryland;	
	/cell_line="HOR"	Web site:	http://www.nisc.nih.gov/
	1..2550	Contact:	nisc_mgc@nhgri.nih.gov
	/gene="HLA-DQB1"	Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastriano, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, I., Young, A., Zhang, L.-H. and Green, E.D.	
	/number=2		
	/functions="antigen presenting molecule"		
	/allele="HLA-DQB1*050401"		
ORIGIN			
Query Match	90.6%; Score 15.4; DB 5; Length 2550;		
Best Local Similarity	94.1%; Pred. No. 1.4e+03;		
Matches	16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1 GCCCACCCTTGCCCGAG 17	FEATURES	Location/Qualifiers
Db	26 GCCCACCCTTGCCCGAG 42	source	1..2623
RESULT 38			/organism="Homo sapiens"
BC018939			/mol_type="mRNA"
LOCUS	BC018939 2623 bp mRNA linear PRI 24-FEB-2004		/db_xref="taxon:9606"
DEFINITION	Homo sapiens cartilage intermediate layer protein 2, mRNA (CDNA clone IMAGE:4126917), partial cds.		/clones="IMAGE:4126917"
ACCESSION	BC018939		/tissue_type="Brain, neuroblastoma"
VERSION	BC018939.2 GI:39644584		/clone_id="NIH MGC 19"
KEYWORDS			/lab_host="DH10B-R"
SOURCE	Homo sapiens (human)		/note="Vector: pOTB7"
ORGANISM	Homo sapiens		<1..2623
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			/gene="CILP2"
1 (bases 1 to 2623)			/note="synonyms: MGC45771, CLIP-2"
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
TITLE		ORIGIN	DIREMSEAAQAQARASGLTRGRVRVQ"
		Query Match	90.6%; Score 15.4; DB 5; Length 2623;
		Best Local Similarity	94.1%; Pred. No. 1.4e+03;
		Matches	16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
		QY	1 GCCCACCCTTGCCCGAG 17

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 20 Row: f Column: 6.

```

Db      1696 GCCCAGCTGGCCCAAG 1712
|||||
BC034926      2623 bp      mRNA      linear      PRI 27-APR-2005
Homo sapiens cartilage intermediate layer protein 2, mRNA (CDNA
clone IMAGE:4124496), partial cds.
ACCESSION    BC034926
VERSION      BC034926.1      GI:22028437
KEYWORDS
SOURCE
ORGANISM     Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE    1 (bases 1 to 2623)
AUTHORS     Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Ditachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
McKernan, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smallus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2623)
NIH MGC Project
Direct Submission
CONSRTM
TITLE       Mammalian Gene Collection Program Team
JOURNAL
PUBMED
REFERENCE   NIH MGC Project
AUTHORS
TITLE       Direct Submission
JOURNAL
REMARK
COMMENT     Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Martin Hirst, Thomas Zeng, Ryan Morin, Michelle Moksa, Johnson
Pang, Diana Mah, Jing Wang, Kieth Fichter, Eric Chuah, Allen
Delaney, Rob Kirkpatrick, Agnes Baross, Sarah Barber, Mabel
Brown-John, Steve S. Chand, William Chow, Ryan Babakaiff, Dave
Wong, Corey Matsuo, Jaclyn Beland, Susan Gibson, Luis delRio, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Mike R. Mayo, Josh Moran, Diana Palmquist, JRC
Santos, Duane Smalish, Jeff Scott, Miranda Tsai, George Yang,
Jacquie Schein, Asim Siddiqui, Steven Jones, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IPAL Place: 43 Row: d Column: 22.
Location/Qualifiers
1. .2623
/organism="Homo sapiens"
/mol_type="mRNA"
FEATURES
source
/db_xref="taxon:9606"
/clone="IMAGE:4124496"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1. .2623
/gene="CILP2"
/note="synonyms: CLIP-2, MGC45771"
/db_xref="GeneID:148113"
<1. .1964
/gene="CILP2"
/codon_start=3
/product="CILP2 protein"
/protein_id="AAH34926.2"
/db_xref="GI:62914018"
/db_xref="GeneID:148113"
/translat="AYQDFTIEVPPSTQRLVVTFFVDPGSEFMDVRVLPFDPGRGAV
YHEVKMRKAPVILHTSQSNTIFLGEDEAPLCELVLPSCAFRRADGKPYSGPVEA
RVTFVDPDLTSAASAPSLRFDVSDGELAPLRTYGMFSVOLRAPGSAEQLQVGPVAV
RVAASQIHMPGHVEALKLNSLPETGLWEESGFRREGSGPRVRERERFVLGNVEI
KERLENDLVPERRCFVKRAYANDKFTPSEQVEGVVTVLLNLEPAPGFSANPRAWG
RFDGAVTGPNGACLPAPCADRDPDAYTALTATLGELEPAPSLRPLPATVGTQF
YLDRLGYRTDHDPAFKRNGFRINLAKPRCPAEANGPVYWRSLRCCQCAPVTAS
HFRFAVEADKYEVNVLPFRGTPTASMTGDLAWNPQEFACFLKVKI OGPQBYMV
RSHNAGSHPTRGQLCLRDARSVRDPERPCTSAACVEFKCSGMLPDQORVDRLTVP
IMPGSCRRVAVNGLLRDYLTRHPPPPAEDPAAFSLAPLDPLCHNYGVTVTTDQSP
RLAKEIAIGRCFDGSSGFSREMKADAGTAVTFQCRPEPPAGRPSPFLQRLLESPTALG
DIRRESEAAQAASGPLTRRRVRQ"
ORIGIN
Query Match      90.6%; Score 15.4; DB 5; Length 2623;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 GCCCAGCTGGCCCAAG 17
|||||
Db      1696 GCCCAGCTGGCCCAAG 1712
RESULT 40
AF420019
LOCUS      AF420019
DEFINITION Aspergillus nidulans 1,4-beta-D-glucan-cellobiohydrolase (cbha)
gene, complete cds.
ACCESSION  AF420019
VERSION     AF420019.1      GI:21449326
KEYWORDS
SOURCE
ORGANISM    Emericella nidulans (anamorph: Aspergillus nidulans)
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
REFERENCE   1 (bases 1 to 3252)
AUTHORS     Lockington, R.A., Rodbourn, L., Barnett, S., Carter, C.J. and
Kelly, J.M.
TITLE       Regulation by carbon and nitrogen sources of a family of cellulases
in Aspergillus nidulans
JOURNAL     Fungal Genet. Biol. 37 (2), 190-196 (2002)
PUBMED
REFERENCE   2 (bases 1 to 3252)
AUTHORS     Lockington, R.A., Rodbourn, L., Barnett, S., Carter, C.C. and
Kelly, J.M.
TITLE       Direct Submission
JOURNAL     Submitted (18-SEP-2001) Molecular Biosciences, University of
Adelaide, North Terrace, Adelaide, SA 5005, Australia
FEATURES
source
1. .3252
/organism="Emericella nidulans"
/mol_type="genomic DNA"
/db_xref="taxon:162425"
/note="anamorph: Aspergillus nidulans"
1231. .2884
/gene="cbha"

```

```
mRNA
join(1231..1498,1555..1938,1995..2884)
/gene="cbha"
/product="1,4-beta-D-glucan-cellobiohydrolase"
join(1283..1498,1555..1938,1995..2735)
/gene="cbha"
/EC_number="3.2.1.91"
/notes="cellulase"
/codon_start=1
/product="1,4-beta-D-glucan-cellobiohydrolase"
/protein_id="AA054069.1"
/db_xref="GI:21449327"
/translation="MYQALLFSALLSVRAQAQAGTAQEVHPSLTHWQCEASGCTE
VAGSVLDSNWRWTHSVVDGTYNCTGNEWDATLCPDNESCAQNCADVADGAYEATYGIT
SNGDSLTLKTSVNSGSRVYLMEDEDTYQFDLLNNEFTFDVDSNPGCLAGALYF
TSMADGGLSKYENTAGAKYGTGCDQCPQDIKFKINGLGNVEGPEPSDSDANAGVG
GMGTCCPEMIWEANSISTAYTTPHPCDSVEOTWCEGDCGTYSDRYGGTCDDGCD
FNSYRWGNTFEPYGAIIDTSKFTVVTQFIADGGSLSSEIKREYVONGEVI PNSESNI
SVGEINSITSEFTQAQTAFGDEIFAQHGLSAMGDAAAMVLLISINDDDHSSMMW
LDSSYPTDADPSQPGVARGTCQAGDPDVVESEHADASVTFNSNFKPGPIGSTF"

ORIGIN
Query Match          90.6%; Score 15.4; DB 4; Length 3252;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACCCTTGCCGAG 17
||||| |||||||
Db 1207 GCCCACCCTTGCCGAG 1223

RESULT 41
CQ727552          3325 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 13486 from Patent WO02068579.
ACCESSION
CQ727552
VERSION
CQ727552.1 GI:42293852
KEYWORDS
.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 13486 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1..3325
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          90.6%; Score 15.4; DB 2; Length 3325;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACCCTTGCCGAG 17
||||| |||||||
Db 3095 GCCCACCCTTGCCCAAG 3111

RESULT 42
AF542080          3602 bp mRNA linear PRI 25-SEP-2002
LOCUS
DEFINITION
Homo sapiens cartilage intermediate layer protein-like protein
C1IP-2 mRNA, complete cds.
ACCESSION
AF542080
VERSION
AF542080.1 GI:23307794
KEYWORDS
.
```

```
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 (bases 1 to 3602)
AUTHORS
Crowl, R.M., Szeto, D. and Farley, D.
TITLE
CILP-2, an isoform of cartilage intermediate layer protein
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 3602)
AUTHORS
Crowl, R.M., Szeto, D. and Farley, D.
TITLE
Direct Submission
JOURNAL
Submitted (29-AUG-2002) Arthritis Research, Novartis
Pharmaceuticals, 556 Morris Ave, Summit, NJ 07901, USA
FEATURES
Location/Qualifiers
1..3602
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="19p13.11"
/chromosome="19"
/tissue_type="osteochondritic cartilage"
54..3524
/notes="glycoprotein; related to cartilage matrix protein
CILP"
/codon_start=1
/product="cartilage intermediate layer protein-like
protein CILP-2"
/protein_id="AAN17826.1"
/db_xref="GI:23307795"
/translation="MASLLPLCLVVAHLAGARDATPEEPMTALGLERRSVYTC
QPSPALDEEASESWFNVDHPGGDGFESLAAIRFYGPARKVCRPRLALERTTD
WALPSAVGERVHLNPTRGFWCLNREQPRRCNSYHVRFCPLSEAGMAGWPGPCSG
SCQPGRLRRRHCPSPAGDAPCPLEAQKCVRCPCGSLDTCEDCHILLGVSVP
SGOPLGARLRDQPGTATSDAHGTFRVPGCADSRANIRMDGFSAGEAQAQAN
GSISVVTIILDKLEKPYLVKHPESVREAGQNTVFCCKASCTPMKKYSWFHNGTLLD
RRAHGGAHLELGLRPDQGIYHCKAWNAGAVRSCTARLTVLAPQCPADCPREY
LILPEDCGQPGSPAYLDVLCPTDTRCPSLAGSSPCGDCASSCCSVRRERREIHC
PGYVLPVKVAECQCKPLPGLVRGVAAADSGEPLRFAILLGQEPGIFTAYQGD
FTIEVPSTQRLVVTVDPSGEFMDAVRVLPFDPRGAGVYHEVKAMRKKAFLVILTSQ
SNIPLEGEDEAPLGBELVLPSPGARADGPKYGPVEARVTVDPDRDLTSAASAPS
RDARSVDRLPRTYGMFSDLRAPGSARQLQVGPVAVRVAASQIHMPGVHVKLW
SLNPETGLWEEESGFRREGSGPRVREERVELGNVEIRERLFDNDVPERRCFVK
VRAYADKFTPSQVEGVVTVLNLEPGPGFSANPRAWGRFDSAVTPNGACLAFCFD
ADPRDAYTALVTATLGELEPAPSLPRPLPATVGTQPYLDRLGYRRTDHDPAFRK
NGRINILAKPRPGDPAEANGVYVWRSLRECQAPVASHFRFARVEADKYEYNNVVF
REGTPASWTGDLAWPNPQBFACFLKVIQGPQYMYRSHNAGSGHPRTRGQLYGL
RDARSVDRLPRTYGMFSDLRAPGSARQLQVGPVAVRVAASQIHMPGVHVKLW
LTRHPPVPAEDPAAFSMLAPLDPLGHNYGVTVTQDQSPRLAKIAIGRCFCGSSDGF
SREMKADAGTAVTFCQREPPAGRPSLFQRLLESFATALGDIRREMSANAQAARASGP
LRTRRGRVRQ"

ORIGIN
Query Match          90.6%; Score 15.4; DB 5; Length 3602;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACCCTTGCCGAG 17
||||| |||||||
Db 3256 GCCCACCCTTGCCCAAG 3272

RESULT 43
AR157945          4183 bp DNA linear PAT 17-OCT-2001
LOCUS
DEFINITION
Sequence 2 from patent US 6251389.
ACCESSION
AR157945
VERSION
AR157945.1 GI:16219890
KEYWORDS
.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 4183)
```



```

AUTHORS      Magna,H., Schaffer,P., Lawton,M., Yocum,S., Mitchell,P.,
              Hutchinson,N. and Murry L.E.
TITLE        Human nucleotide pyrophosphohydrolase-2
JOURNAL      Patent: US 6251389-A 2 26-JUN-2001;
FEATURES     Location/Qualifiers
              1..4183
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN
Query Match      90.6%; Score 15.4; DB 2; Length 4183;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCCGAG 17
    |||||
Db 3254 GCCCAGCTTGCCCAAG 3270

RESULT 44
BD103585
LOCUS      BD103585                4183 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Human nucleotide pyrophosphohydrolase-2.
ACCESSION  BD103585
VERSION     BD103585.1 GI:22649159
KEYWORDS    JP 2001526886-A/1.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 4183)
            Magna,H., Schaffer,P., Lawton,M., Yocum,S.A., Mitchell,P.G.,
            Hutchinson,N. and Murry,L.E.
            Human nucleotide pyrophosphohydrolase-2
            Patent: JP 2001526886-A 1 25-DEC-2001;
            INCYTE PHARMACEUTICALS INC
COMMENT     OS Homo sapiens (human)
            PN JP 2001526886-A/1
            PD 25-DEC-2001
            PF 02-DEC-1998 JP 2000525259
            PR 22-DEC-1997 US 08/996083
            PI HOLLY MAGNA,PAUL SCHAFFER,MICHAEL LAWTON,SUE A YOCUM,PETER G
            MITCHELL,
            PI NANCY HUTCHINSON,LYNN E MURRY
            PC C12N9/16,A61K38/46,A61K39/395,A61K45/00,A61P19/02,A61P35/00,
            PC C07K16/40,C12N15/09,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
            A61K37/54,
            PC C12N15/00
            CC a or g or c or t, unknown, or other
            CC 1388013
            FH Key      Location/Qualifiers
            FT unsure   4180.
FEATURES     source
              1..4183
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"

ORIGIN
Query Match      90.6%; Score 15.4; DB 2; Length 4183;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCCGAG 17
    |||||
Db 3254 GCCCAGCTTGCCCAAG 3270

RESULT 45
CQ850573
LOCUS      CQ850573                4232 bp      DNA      linear      PAT 23-AUG-2004
DEFINITION Homo sapiens cDNA FLJ45835 fis, clone NT2R9808057, weakly similar
            to Homo sapiens cartilage intermediate layer protein, nucleotide
            pyrophosphohydrolase (CILP).
ACCESSION  AK127735
VERSION     AK127735.1 GI:34534774
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
            Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
            Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 4232)
            Isogai,T. and Yamamoto,J.
            Submitted (15-JUL-2003) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail: flj-cdna@nifty.com, Tel.81-438-52-3975, Fax:81-438-52-3986)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI); (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: Reverse Proteomics Research Institute, HRI and
            RAB.
FEATURES     source
              1..4232
              Location/Qualifiers

```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP0008057"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RFP8"
/notes="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor
retinoic acid (RA) induction."

ORIGIN

Query Match      90.6%; Score 15.4; DB 5; Length 4232;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
Db 3323 GCCCAGCGCTGGCCCAAG 3339

RESULT 47
LOCUS AB209580 4641 bp mRNA linear PRI 31-MAR-2005
DEFINITION Homo sapiens mRNA for MHC class II antigen variant protein.
ACCESSION AB209580
VERSION AB209580.1 GI:62088739
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
Ohara,O., Nagase,T. and F.Kikuno,R.
TITLE None Title
JOURNAL Published Only in Database (2005)
REFERENCE 2 (bases 1 to 4641)
AUTHORS Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
Ohara,O., Nagase,T. and F.Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail:cdnaif@kazusa.or.jp,
URL:http://protein.gsc.riken.go.jp/, Tel:81-438-52-3330,
Fax:81-438-52-3931)
COMMENT This work was supported in part by the National Project on Protein
Structural and Functional Analysis, Ministry of Education, Culture
, Sports, Science and Technology of Japan. Totoki Y, Toyoda A,
Takeda T, Sakaki Y, Tanaka A, Yokoyama S. RIKEN Genomic Sciences
Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
e-mail: aktanaka@postman.riken.go.jp
URL: http://protein.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
1..4641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="sj02456"
/tissue_type="spleen"
/notes="this clone is also named as hek003001744
vector:pBluescriptII SK plus"
<1..4641
/gene="MHC class II antigen variant"
<1..499
/gene="MHC class II antigen variant"
/inference="non-experimental evidence, no additional
details recorded"
/notes="Start codon is not identified."
/codon_start=2
/product="MHC class II antigen variant"
/protein_id="BAD92817.1"

gene
CDS
```

```
/db_xref="GI:62088740"
/translation="TTTAFEFVSMWKKALRIPGDLRVATVTLMAMLSLLAEGRD
SPDEFQFKMGCMYFTNGRLVRIYINREYARFSDVSDGVYRAVTPQGRPDAY
WNSQKEVLEGTRAEIDTVCRHNYEAVFRGILQRRGRRRPSVRPTLGRDPEFLCREG
WGRGL"

ORIGIN

Query Match      90.6%; Score 15.4; DB 5; Length 4641;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
Db 436 GCCCAGCGCTGGCCGAG 452

RESULT 48
LOCUS AY375870 5425 bp DNA linear PRI 31-AUG-2004
DEFINITION Homo sapiens 4A0H 100040 WAL FD MHC class II antigen (HLA-DQB1)
gene, HLA-DQB1*0602 allele, exons 1 through 4, and partial cds.
ACCESSION AY375870
VERSION AY375870.1 GI:37021370
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 5425)
AUTHORS Wood,J.M., Simons,M.J. and Ashdown,M.L.
TITLE Comparative haplotype analysis of polymorphism within the DQB1
gene. Continuous sequencing from promoter to exon 4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5425)
AUTHORS Wood,J.M., Simons,M.J. and Ashdown,M.L.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2003) Genetype Pty Ltd, 60-66 Hanover Street,
Fitzroy, Victoria 3065, Australia
FEATURES
Location/Qualifiers
1..5425
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/haplotype="DQB1*0602 DQAL*01021 DRB1*15011"
/cell_line="4A0H 100040 WAL FD"
<236..>5425
/gene="HLA-DQB1"
/allele="HLA-DQB1*0602"
join(<236..344,1776..2045,4595..4876,5386..>5425)
/gene="HLA-DQB1"
/product="MHC class II antigen"
join(236..344,1776..2045,4595..4876,5386..>5425)
/gene="HLA-DQB1"
/allele="HLA-DQB1*0602"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAQ88044.1"
/db_xref="GI:37021371"
/translation="MSWKALRIPGDLRVATVTLMAMLSLLAEGRSPDFVFPQK
GMCYFTNGRLVRIYINREYARFSDVSDGVYRAVTPQGRPDAYWNSQKEVLEG
TRAEIDTVCRHNYEAVFRGILQRRVPTVTSRKTALNHNLLVCSVTDFYQGIK
VRFRNDQEEETAGVYVSTPLIRNGDWTTFQILVMLEMTFQRGDVYTCVHEHPSLQSPITV
EWRAQSSAQSKMLSG"
<236..344
/gene="HLA-DQB1"
/allele="HLA-DQB1*0602"
/number=1
1776..2045
/gene="HLA-DQB1"
/allele="HLA-DQB1*0602"
/number=2

gene
mRNA
CDS
exon
exon
```

```

exon      4595..4876
           /gene="HLA-DQB1"
           /allele="HLA-DQB1*0602"
           /number=3
exon      5386..55425
           /gene="HLA-DQB1"
           /allele="HLA-DQB1*0602"
           /number=4
ORIGIN
Query Match      90.6%; Score 15.4; DB 5; Length 5425;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1  GCCCAGCTTGGCCGAG 17
        ||||| ||||| |||||
Db      2071 GCCCACCCTTGGCCGAG 2087

RESULT 49
LOCUS      AY375871      5599 bp      DNA      linear      PRI 31-AUG-2004
DEFINITION Homo sapiens 10IHS 9010 AMAI MHC class II antigen (HLA-DQB1) gene,
            HLA-DQB1*0602 allele, exons 1 through 4, and partial cds.
ACCESSION      AY375871
VERSION      AY375871.1 GI:37021384
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE      1 (bases 1 to 5599)
AUTHORS      Wood, J.M., Simons, M.J. and Ashdown, M.L.
TITLE      Comparative haplotype analysis of polymorphism within the DQB1
            gene. Continuous sequencing from promoter to exon 4
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 5599)
AUTHORS      Wood, J.M., Simons, M.J. and Ashdown, M.L.
TITLE      Direct Submission
JOURNAL      Submitted (25-AUG-2003) Genetype Pty Ltd, 60-66 Hanover Street,
            Fitzroy, Victoria 3065, Australia
FEATURES
            source
            1..5599
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /haplotypes="DQB1*0602 DQA1*01021 DRB1*1503"
               /cell_line="10IHS 9010 AMAI"
               <410..>5599
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*0602"
               join(<410..518,1950..2219,4769..5050,5560..>5599)
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*0602"
               /product="MHC class II antigen"
               join(<410..518,1950..2219,4769..5050,5560..>5599)
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*0602"
               /codon_start=1
               /product="MHC class II antigen"
               /protein_id="AAQ88045.1"
               /db_xref="GI:37021385"
               /translation="MSWKKALKIPGDLRVATVTLMLSLLAEGRDSPEDFVYQFK
               GMCYFTNGTERVLRYIYNREYAFDSVGVYRAVTPQGRPDAYNWSQKEVLEG
               TRALDITVCRHNYEAFRGILQRVEPTVTISPSRTEALNHNLLVCSTDFYPCQIK
               VWRFRNDOEETAGVSTPLIRNGDWTFOILVLEMTFQGRDVTTCVHVEHPSLQSPITY
               EWRAQSESAQSKMLSG"
               <410..518
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*0602"
               /number=1
               1950..2219
exon
exon
```

```

/allele="HLA-DQB1*0603"
/number=1
1950. .2219
/gene="HLA-DQB1"
/allele="HLA-DQB1*0603"
/number=2
1950. .2219
/gene="HLA-DQB1"
/allele="HLA-DQB1*0603"
/number=3
4770. .5051
/gene="HLA-DQB1"
/allele="HLA-DQB1*0603"
/number=4
5561. .5600
/gene="HLA-DQB1"
/allele="HLA-DQB1*0603"
/number=4

ORIGIN
Query Match          90.6%; Score 15.4; DB 5; Length 5600;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
||||| ||||| ||||| |||||
Db 2245 GCCCAGCGTTGGCCGAG 2261

RESULT 51
AY375873
LOCUS      5600 bp DNA linear PRI 31-AUG-2004
DEFINITION Homo sapiens 4A0H 100070 EMJ MHC class II antigen (HLA-DQB1) gene,
            HLA-DQB1*06041 allele, exons 1 through 4, and partial cds.
ACCESSION  AY375873
VERSION     AY375873.1 GI:37021414
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE   1 (bases 1 to 5600)
AUTHORS    Wood,J.M., Simons,M.J. and Ashdown,M.L.
TITLE      Comparative haplotype analysis of polymorphism within the DQB1
            gene. Continuous sequencing from promoter to exon 4
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 5600)
AUTHORS    Wood,J.M., Simons,M.J. and Ashdown,M.L.
TITLE      Direct Submission
JOURNAL    Submitted (25-AUG-2003) Genetype Pty Ltd, 60-66 Hanover Street,
            Fitzroy, Victoria 3065, Australia
FEATURES   source
            1. .5600
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /haplo_type="DQB1*06041 DQAI*01021 DRB1*1302"
               /cell_lines="4A0H 100070 EMJ; 101HWS 9097 EMJ"
               <410. .>5600
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*06041"
               join(<410. .518,1950. .2219,4770. .5051,5561. .>5600)
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*06041"
               /product="MHC class II antigen"
               join(410. .518,1950. .2219,4770. .5051,5561. .>5600)
               /gene="HLA-DQB1"
               /allele="HLA-DQB1*06041"
               /codon_start=1
               /product="MHC class II antigen"
               /protein_id="AAQ88047.1"
               /db_xref="GI:37021415"
               /db_xref="GI:37021415"
               TRAEADTVCRHNYEAFRGILQORVEPTVTISPSRTEALNHNLLVCSTVDYFPGQIK
               VWRFRNDEETAGVSVSTPLRNGDWTQTQLVWLEMTQRGDVTYCHVEHRSQSPITV

```

```

EWRAQSESAQSKMLSG"
<410. .518
/gene="HLA-DQB1"
/allele="HLA-DQB1*06041"
/number=1
1950. .2219
/gene="HLA-DQB1"
/allele="HLA-DQB1*06041"
/number=2
4770. .5051
/gene="HLA-DQB1"
/allele="HLA-DQB1*06041"
/number=3
5561. .5600
/gene="HLA-DQB1"
/allele="HLA-DQB1*06041"
/number=4

ORIGIN
Query Match          90.6%; Score 15.4; DB 5; Length 5600;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
||||| ||||| ||||| |||||
Db 2245 GCCCAGCGTTGGCCGAG 2261

RESULT 52
BLVCG/c
LOCUS      BLVCG      8714 bp ss-RNA linear VRL 28-APR-1993
DEFINITION Bovine leukemia virus (proviral), complete genome.
ACCESSION  K02120
VERSION     K02120.1 GI:210767
KEYWORDS    complete genome; endonuclease; env protein; gag protein; pol
            protein; proviral gene; reverse transcriptase; surface
            glycoprotein; transmembrane protein.
SOURCE      Bovine leukemia virus
ORGANISM    Bovine leukemia virus
            Viruses; Retro-transcribing viruses; Retroviridae;
            Orthoretrovirinae; Deltaretrovirus.
            1 (bases 1 to 8714)
            Sagata,N., Yasunaga,T., Tsuzuku-Kawamura,J., Ohishi,K., Ogawa,Y.
            and Ikawa,Y.
            Complete nucleotide sequence of the genome of bovine leukemia
            virus: its evolutionary relationship to other retroviruses
            Proc. Natl. Acad. Sci. U.S.A. 82 (3), 677-681 (1985)
            2983308
            Original source text: Bovine leukemia virus RNA.
            The organization of the BLV genome is as follows: 5' LTR - gag -
            pol - env - pX-BL - 3' LTR, where pX-BL is a region containing
            several open reading frames.
            [1] suggests, based on the pattern of conservation in the pol
            region, that BLV and human T-cell leukemia virus (HTLV) constitute
            a separate group, designated type E, of Oncovirinae. The initiation
            of translation of the pol reading frame is undetermined. A
            potential splice acceptor (at 4391-4401) for the env mRNA is
            identified by similarity to the consensus sequence [1].
            Location/Qualifiers
            1. .8714
               /organism="Bovine leukemia virus"
               /mol_type="genomic RNA"
               /db_xref="taxon:11901"
               1. .530
               /standard_name="5' LTR"
               628. .1806
               /gene="gag"
               628. .1806
               /gene="gag"
               /standard_name="gag-Pr45"
               /codon_start=1
               /protein_id="AAA42784.1"
               /db_xref="GI:210768"

FEATURES   source
            1. .8714
               /organism="Bovine leukemia virus"
               /mol_type="genomic RNA"
               /db_xref="taxon:11901"
               1. .530
               /standard_name="5' LTR"
               628. .1806
               /gene="gag"
               628. .1806
               /gene="gag"
               /standard_name="gag-Pr45"
               /codon_start=1
               /protein_id="AAA42784.1"
               /db_xref="GI:210768"

```

```
/translation="MGNPSYNPPAGISPSDWLNLLOSAQRINPRPSPDFTDLKNYI
YHFKTKQKPTFTSGGPTSCPPGRFVPLVLAITNEVLSNEGAGCAGSAPPEOPP
YDPAILPIISEGNRHRWALRELQDIKEIENKARQSOWIQTLEAILQADPTP
ADLEGQCIYASPVDTAHTSLTAIAAENATSRVLTKTGLTQOQAPNAGDL
RSQYNLWLAGKISLLVLQGPWSTIVOGPAESSVEFVRNKLQISLADNLPGVLRNP
LTLPLWMLTSEKSCFGRGASGGAKTAGLRTIGPPRMKQPALVHTTFCPKMGP
RQAPAPRPPEPGPCVCLKEGHWARDCTKATGPPPGPCICKDPSHWKRDCTPLKSN
"
```

mat_peptide

628. .954

/gene="gag"

/product="gag-p15 polyprotein"

955. .1545

/gene="gag"

/product="gag-p24 polyprotein"

1597. .1803

/gene="gag"

/product="gag-12 polyprotein"

2467. .4875

/gene="pol"

/codon_start=1

/product="pol protein"

/protein_id="AAA42785.1"

/db_xref="GI:210769"

```
/translation="VHDLRATNALTKPIPALSPGPPDLTAIPTHPHIICLDLKDAAF
QIPVEDRFRVLSFTLPSGGLQHRREARFALPQGFINSALPERALQERLQVSA
FSQSLVSYMDILYASTERQSCYOALARNLDLQGVASERTSQTPSPVFLQG
MVRQIVTYSQTLQISSLHQLQAVLDQVMSRGTTTPRRLQLLYSLKRRH
DPAIQLSPQLOAGIALRQALSHNRSYNEQELLYVHLTRAGSLVLFQKGAQ
FPLAFOTPLDQNASPWGLLLGLCOYLQALSSYAKPILKYHNLPTKSLDNWIO
SSEPRVQLQLQPOISSOGIOPPGPKTLITRAEVFLTPQFSPDPTPAALCLFSDG
ATGACGLMDHLLDFOAVAPESAKOGELAGLAGLAAPPEPVNIWDSKLYLSL
LRTLVGLAWLPDPVPSVALYKLSLRHPALVGVHRSASPIASLANNVVDQLLP
LETPEQMKLTHCNSLRWPNRISPRISATCETCOKLNKGGKMTIQRG
WAPNHIOADITHYKIOFTYALHVFVDYTSAGTHASAKRGLTTQTTLEGLEAIVHL
GRPKNLDOQANTYTKTFVFCQFQVSLSHVYPNPTSSGLDERTNGLKLLLSKY
HLDEPLPMTQALSRALWTHQINLLPIKTRWELHSPPLAVISEGGETPKGSDKLF
LYLLPGQNNRRWLGLPALVEASGALLATDPVWVPHRLKAKFKCLKNDGPDAHNR
SSDG"
```

gene

4821. .6368

/gene="env"

4821. .6368

/gene="env"

/codon_start=1

/product="envelope polyprotein"

/protein_id="AAA42786.1"

/db_xref="GI:210770"

```
/translation="MPKERRRRPQPIIRWVSLTLTLALCRPIQTWRCSLSLGNQO
WMTAYNOEKFISIDQILEAHNOSPFCASPRVTLDSVNGVPKLYWPPGRRRFGA
RAMVYDCEPCPVYGDRCFDCHWDNASQDQSFVYNHQLFLHLKQCHGIFLTW
EINGYDLITSLHKIDPPDPQFQLNSDWVPSVRWALLNQATAPFCAICWEP
SPWAPPELLVYNTKISSGGGLAPDQAFIWNSSSNFTQGMHPHSPORLLFNVSQGN
ALLPPLSLVNLASAPPRVRSVPAALTLGALSGLGTGINVASLSHORLTS
LIHVLEQDQIIRGLDPLQSRVSDWQWPNNDLGTAWVRETIHSLVLSLFLALEFL
FLAPCLIKLTSRLKLLRQAPFPEISLTPKPSDYLQALLPSAPEIYSHLSVPKPDY
INLRCP"
```

sig_peptide

4821. .4919

/gene="env"

4920. .5723

/gene="env"

/product="gp51 surface glycoprotein (env)"

5724. .6365

/gene="env"

/product="gp30 transmembrane protein (env)"

8185. .8714

/standard_name="3' LTR"

/standard_name="3' LTR"

ORIGIN

63 bp upstream of PrU1 site.

Query Match 90.6%; Score 15.4; DB 10; Length 8714;
Best Local Similarity 94.1%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 GCCCAGCGCTTGGCCGAG 17
      |||||
Db 4288 GCCCAGCGCTTGGCCGAG 4272
```

RESULT 53

AE004474/c

LOCUS

DEFINITION

Pseudomonas aeruginosa PAO1, section 35 of 529 of the complete genome.

ACCESSION

AE004474

VERSION

AE004474.1

KEYWORDS

SOURCE

ORGANISM

Pseudomonas aeruginosa PAO1

Pseudomonas aeruginosa PAO1

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE

1 (bases 1 to 10694)

AUTHORS

Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z. and Paulsen, I.T.

TITLE

Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen

JOURNAL

NATURE 406 (6799), 959-964 (2000)

PUBMED

10984043

REFERENCE

2 (bases 1 to 10694)

AUTHORS

Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.

TITLE

Direct Submission

JOURNAL

Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA

REFERENCE

3 (bases 1 to 10694)

AUTHORS

Pseudomonas aeruginosa Community Annotation Project (PseudocAP)

CONSRMT

Direct Submission

TITLE

Submitted (04-FEB-2003) Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada

JOURNAL

COMMENT

This represents the February 3, 2003 version of the continually updated, reviewed, Pseudomonas aeruginosa PAO1 genome annotation, from PseudocAP (see <http://www.pseudomonas.com> for latest updates and links to alternate annotations). PseudocAP is coordinated by Fiona S.L. Brinkman (Simon Fraser University, Canada) and Robert E.W. Hancock (University of British Columbia, Canada). We welcome submission through www.pseudomonas.com of any proposed changes.

'Protein name confidence' is used to rate our confidence of the accuracy of the protein name.

Class 1: Function experimentally demonstrated in P. aeruginosa.

Class 2: Function of highly similar gene experimentally demonstrated in another organism (and gene context consistent in terms of pathways its involved in, if known).

Class 3: Function proposed based on presence of conserved amino acid motif, structural feature or limited sequence similarity to an experimentally studied gene.

Class 4: Homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences.

Location/Qualifiers

1. .10694

/organism="Pseudomonas aeruginosa PAO1"

/mol_type="genomic DNA"

FEATURES

source

```

/strain="PA01"
/db_xref="taxon:208964"
166..1152
/locus_tag="PA0360"
166..1152
/locus_tag="PA0360"
/notes="Protein name confidence: Class 4 (homologs of
previously reported genes of unknown function, or no
similarity to any previously reported sequences)"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG03749.1"
/db_xref="GI:9946211"
/transl_table=11
/translation="MFSQWFLSVLALVPSFAWAEDAVSDLAREQSKAEQGGKRGDE
LFAEFTLRQAETGEPEQALFGAYFYQQNTYASAREWKKAAAGMVAQQLAML
YRGDGGQDKTEAARWFKAAEGDAAGQNAEMGVLYRGGEVDQDRVKAGTWFERAA
AGSEDAETNLGHFYLDSSQGVYSATSDDEEAALLDRYAGSREKAFOWECKANTQGDAR
AQFKVGAYWNGSGAGWNKQARLWLSKAAQQQDADAIWLSESAANAATWTRLENWTH
AALDGEUSGVADCSGAADRRHSGRMEGAPSDPEAVEASAEQAQVAEASQ"
/locus_tag="PA0361"
/complement(1257..2990)
/locus_tag="PA0361"
/complement(1257..2990)
/locus_tag="PA0361"
/notes="Protein name confidence: Class 3 (function proposed
based on presence of conserved amino acid motif,
structural feature or limited sequence similarity to an
experimentally studied gene); Subcellular localization:
Periplasmic; Subcellular localization confidence: Class 2"
/codon_start=1
/transl_table=11
/product="probable gamma-glutamyltranspeptidase precursor"
/protein_id="AAG03750.1"
/db_xref="GI:9946212"
/transl_table=11
/translation="MSFLNLPVRLTVAACGLSLLSLAHADSPRGDAHAPQAAVATP
HPAATVAGLTLANGNAFDAAAIAALAAVAEPYSGGLGGGPELLRQAQOPTYFF
LDARERAPKAAVDYMRNRNGVDPRLSDVGLPAAAIPLPALVELSRVGRKPLADN
LVPAIRLAVGVDRYDRRAAMRLKRPETARIFLDKGGIPDENLLRQPLA
RTLRLGRYGRIGYEGEATDKLLTGVRAGGIMSAADLRDYRIERRPLEVKANGR
ELISAPPPSAGGVALGALQMLEQLPWQKAEFPVQARHYVLEVLRRAYDRGLGLDDPF
VANPLSLAPDYLRKLAAGIDPRRATPSSALPEAPAWREGDHTTHFAVIDDAQNVA
ATSVNLPGCAFTVPTDGVNLNEMDDFADTQGANSGYGLAGSQAANAAGKRPGLS
MSPFLSPDTPFAAFGTGGSRIFSMVLINLQYFDGRPVQWVAAPRHHQYKPDV
EYEPRAFDAAEALRRRGYQLKRLERSYGNQQLYLDKRSRVDAAASDRFGIGTSL
SLEAQAPRPH"
/complement(3118..3369)
/locus_tag="PA0362"
/complement(3118..3369)
/locus_tag="PA0362"
/notes="Protein name confidence: Class 2 (high similarity
to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="ferredoxin [4Fe-4S]"
/protein_id="AAG03751.1"
/db_xref="GI:9946213"
/translation="MSLKITDDCTNCVCEPCNGAISQGEIYVIDPNLCTECVGH
YDEPQQCVQVDCIPLDANVESKDQLMEKYKITGKA"
/complement(3490..3969)
/locus_tag="PA0363"
/locus_tag="PA0363"
/notes="Protein name confidence: Class 2 (high similarity
to functionally studied protein); Alt protein name: PPA1"
/codon_start=1
/transl_table=11
/product="phosphopantetheine adenylyltransferase"
/protein_id="AAG03752.1"

```

```

/db_xref="GI:9946214"
/translation="MNRVLYPGTFDPTIKGHGDLIERASRLFDHVIIVAAASPKKNPL
FSLEQRVALAQEVTKHLPNVEVVGFTLLAHFVKQKCANVFLRLRAVSDFEYFQLA
NMNRQAPDVESMFLTPSEKYSFISSTLVEIRIAALGGDI SKFVHPVADALAEFRK"
/complement(4121..5716)
/locus_tag="PA0364"
/complement(4121..5716)
/locus_tag="PA0364"
/notes="Protein name confidence: Class 3 (function proposed
based on presence of conserved amino acid motif,
structural feature or limited sequence similarity to an
experimentally studied gene)"
/codon_start=1
/transl_table=11
/product="probable oxidoreductase"
/protein_id="AAG03753.1"
/db_xref="GI:9946215"
/translation="MPVPDLFAEGLARGWKTYNGSRLLDDLEADVVVVGTAGGGT
TAEILSAGAKVLLVEGPKLTSDFKQWQADAYPELYQEGIGRMSKDGATITLQGRA
VGGTTLINWTSSTPTEPTLQHWAOYGVGVKHSADMAPFERKEQRLHVPALPEN
ANNVDVRLGCEKLGHWKVIIPRNVLGCWNLGYCGMCPVNAKQSMVLTITPATLDKGG
ELLYLARANRLLDGDKVTLECLGMDERCVPANRRIRVRARHYVLSGGGINTPAIL
LRKAPDPQSQRVGKRTPLHTWFSAGLFDVINPFYGAQSIYSDHFWMDGQVGTGRMS
YKLEVPDLPQSLASVLLGGFSGDNALRMEQLPHTNMMLALLRDCDFHDPDSARGSVLGRG
DDTPVDYRDTYTDGIRAFHTMAEIQFAAGAREVPYVHSDARSANKLQEAAGIID
GLRIEYTRTGSAAHVMGCGAMEDPAVAVADSLGRHHLNLSIHGDSLPTTSIGAN
PQUSVGLTQALASALAERLARA"
/complement(5772..6320)
/locus_tag="PA0365"
/complement(5772..6320)
/locus_tag="PA0365"
/notes="Protein name confidence: Class 4 (homologs of
previously reported genes of unknown function, or no
similarity to any previously reported sequences)"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG03754.1"
/db_xref="GI:9946216"
/translation="MSDTTLESAGLSRRSLMKVGLIGAFATAGVTASLTGCSAEKP
ASGLEKVRSDLPFLRALLPVLMLGAVSAEQMPKAVEGAIQSLDHNLARLSPKFLT
QQLFDVLALPTRGPLTGIWGSWENASGDDVRAFLSRWENSFIGLLRMGHSSLMQLAM
MAMVAREANAHGCGYGPCKIA"
/complement(6375..7805)
/locus_tag="PA0366"
/complement(6375..7805)
/locus_tag="PA0366"
/notes="Protein name confidence: Class 3 (function proposed
based on presence of conserved amino acid motif,
structural feature or limited sequence similarity to an
experimentally studied gene)"
/codon_start=1
/transl_table=11
/product="probable aldehyde dehydrogenase"
/protein_id="AAG03755.1"
/db_xref="GI:9946217"
/translation="MVADIAYLQSQSQEINQLEALFAQRQAAVLAQPMPDATQVQWL
KALRDLFFKQALIEADIRDVSARSADETLLAIMPSLGHGIAAKRVKWKMPARR
AVGLQFQPAQAQVYVQPLVGVYVWPNYPLFLSIGPLTGAAGNRVMIKXSESTPA
TGRLLKOLLARIFFEDQVAVVLGVGVDSKLPFDHLLFTGATSVGHVWRAAEFV
LTPVTLELGGKSPAIIVSDSPVPMKDAERIAFGKSLNAGQTCVADPYLVPSRRVEFV
SQYKEVVGQFPFRLSDNPDYTAIINERQLGRGLRGLDDAREKGATLVFLFAEQQRRL"

```

Query Match 90.6%; Score 15.4; DB 15; Length 10694;

Best Local Similarity 94.1%; Pred. NO. 1.6e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTCCCGAG 17

Db 7832 GCCCAGCTTCCCGAG 7816

RESULT 54

```
AB168089
LOCUS AB168089 16651 bp DNA linear PLN 08-JUN-2004
DEFINITION Chlamydomonas reinhardtii Lcr1 gene for low-CO2 inducible Myb
transcription factor LCR1, complete cds.
ACCESSION AB168089
VERSION AB168089.1 GI:46020169
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
1
Kucho,K., Yoshioka,S., Taniguchi,F., Ohyama,K. and Fukuzawa,H.
Cis-acting elements and DNA-binding proteins involved in
CO2-responsive transcriptional activation of Cahl encoding a
periplasmic carbonic anhydrase in Chlamydomonas reinhardtii
Plant Physiol. 133 (2), 783-793 (2003)
2
Yoshioka,S., Taniguchi,F., Miura,K., Inoue,T., Yamano,T. and
Fukuzawa,H.
The Novel Myb Transcription Factor LCR1 Regulates the
CO2-Responsive Gene Cahl, Encoding a Periplasmic Carbonic Anhydrase
in Chlamydomonas reinhardtii
Plant Cell 16 (6), 1466-1477 (2004)
3
15155888
3 (bases 1 to 16651)
Fukuzawa,H., Yoshioka,S. and Yamano,T.
Direct Submission
Submitted (18-MAR-2004) Hideya Fukuzawa, Kyoto University, Division
of Integrated Life Science, Graduate School of Biostudies,
Sakyo-Ku, Kitashirakawa Oiwake-cho, Kyoto, Kyoto 606-8502, Japan
(E-mail:fukuzawa@lif.kyoto-u.ac.jp, Tel:81-75-753-6391,
Fax:81-75-753-6127)
LOCATION/Qualifiers
1. 16651
/organism="Chlamydomonas reinhardtii"
/mol_type="genomic DNA"
/db_xref="taxon:3055"
/cell_lines="C85-20"
2454..15908
/feature
misc_feature
5318..10420
/feature
misc_feature
5934..5940
/enhancer
/standard_name="enhancer element consensus (EEC)"
/inference="non-experimental evidence, no additional
details recorded"
/citation=[1]
join(6457..6548,6668..6705,6938..6965,7089..7528,
7731..8941)
/gene="Lcr1"
join(6457..6548,6668..6705,6938..6965,7089..7528,
7731..8941)
/gene="Lcr1"
/codon_start=1
/product="low-CO2 inducible Myb transcription factor LCR1"
/protein_id="BAD13491.1"
/db_xref="GI:46020170"
/translat="MTETDHRRSRDPWSRAOSLRILQHLVGLGNSWTEIAKOLPQRTQ
NDCKNFFGALRAKRYGNDLVYARALPPASACGSGWEQKRGDPALTRAAYKA
AMQVAAEVAEQHEKQSRQSQEGEDGCGSGAGATRAEDGEGPAGVAASRRSSSV
SVGDGAAPTADGDMDTQEDAAAPACAPASAAAPSPVPGDVSVRRSSSTGDTVVVTD
AGTRTVAAAGVAGWRSVAAAAMPAPAAVAVVMPVPVPAASVVAASGVLGAAVPA
AGAPDRLSLQSLQPPPHGFALPQSAAPALGSSASPFWQHQHLMGPRVOLLSH
ESLALHQHOQAQSHSHVHLVAPPFLOQHONPHOHLMVELEGAGAFOLQHOHL
HPHVQSGSLPTAAAPSCSNAPAPAPAGAAAGLPPAPAPAPAPALVCPCT
SAACAAATGVPAPASVRCCTCOCRCICCRSCGGRGLRLLPSPAAAAAACCC
RCLCLRLRAARRSACRRALAGTRRLRSVPVQRLQPGGDRHHHPHHNHCRCGCC
RCCWRRSGCQDRARLGRGGHGLWGPAAATEGACWR"
8942..10125
3'UTR
```

```
polyA signal 10108..10112
polyA_site 10125
ORIGIN
Query Match 90.6%; Score 15.4; DB 4; Length 16651;
Best Local Similarity 94.1%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCTTGGCCGAG 17
Db 4417 GCCCACCTTGGCCGAG 4433
RESULT 55
BX000454
LOCUS BX000454 59316 bp DNA linear ROD 17-JUN-2004
DEFINITION Mouse DNA sequence from clone RP24-276N4 on chromosome 4, complete
sequence.
ACCESSION BX000454
VERSION BX000454.33 GI:48927591
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 59316)
Palmer,S.
REFERENCE
AUTHORS Palmer,S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 18, 2004 this sequence version replaced gi:48639588.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP24-276N4 is
from a Male (C57BL/6J) mouse BAC Library VECTOR: pTARBAC1.
FEATURES
Location/Qualifiers
1. 59316
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP24-276N4"
/clone_lib="RPC1-24"
ORIGIN
```


Query Match 90.6%, Score 15.4; DB 6; Length 59316;
Best Local Similarity 94.1%; Pred. No. 2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTGGCCGAG 17
|||||
Db 38622 GCCCAGCGCTGGCCGAG 38638

RESULT 56
AC165982
LOCUS AC165982 Bos taurus clone CH240-173K8, *** SEQUENCING IN PROGRESS ***, 20
DEFINITION unsorted pieces.
ACCESSION AC165982.1 GI:71037423
VERSION HTG; HTGS PHASE1
KEYWORDS Bos taurus (cattle)
SOURCE Bos taurus
ORGANISM Bos taurus

REFERENCE 1
AUTHORS Muzny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P., Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R., Beraducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P., Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I., Chacko,J., Chahrouh,M., Chavez,D., Chen,A., Chen,R., Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A., Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Delgado,O., Denson,S., Derman,C., Ding,Y., Dinh,H., Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K., Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M., Fernandez,S., Fernandez,P., Flagg,N., Forbes,L., Foster,P., Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T., Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D., Gonzalez-Garay,W., Guevara,W., Holder,M., Holland,W., Haerberlen,K., Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P., Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J., Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B., Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S., Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F., Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B., Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D., London,P., Lopez,J., Lorensuhewa,L., Lozado,R., Luk,T., Madu,R., Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E., McClelland,H., McPherson,J., Mercadao,C., Metzker,M., Mlivosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M., Murray,D., Nazareth,L., Ngo,D., Nguyen,N., Norwig-Eastaugh,E., Nott,A., Nwaelelemeh,O., O'Brien,M., Ochi-Okorie,C., Odeh,E., Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B., Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T., Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J., Rabata,D., Rachin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S., Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W., Santibanez,J., Santos,R., Savary,G., Scherer,S., Shen,H., Shen,Y., Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Statok,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R., Trejos,Z., Umani,K., Vargo,C., Verduzco,D., Villasana,D., Virk,D., Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J., Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R., Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L., Zhang,Z., Zhou,J., Weinstein,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 56661)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT ----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help.bcm.tmc.edu
----- Project Information
Center project name: FKEV
Center clone name: CH240-173K8
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 68288 bases at least Q40
Consensus quality: 73107 bases at least Q20
Estimated insert size: 94571; sum-of-contigs estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3787: contig of 3787 bp in length
* 3788: gap of unknown length
* 6174: contig of 2287 bp in length
* 6274: gap of unknown length
* 6275: contig of 3914 bp in length
* 10188: contig of 3914 bp in length
* 10189: gap of unknown length
* 10288: contig of 2080 bp in length
* 12369: gap of unknown length
* 12369: gap of unknown length
* 12468: contig of 4812 bp in length
* 12469: gap of unknown length
* 17280: contig of 2833 bp in length
* 17381: gap of unknown length
* 20214: contig of 5313 bp in length
* 20313: gap of unknown length
* 25626: contig of 5313 bp in length
* 25627: gap of unknown length
* 25727: contig of 2148 bp in length
* 27874: gap of unknown length
* 27875: contig of 2280 bp in length
* 27975: gap of unknown length
* 30255: contig of 2937 bp in length
* 30355: gap of unknown length
* 33291: contig of 5436 bp in length
* 33392: gap of unknown length
* 38828: contig of 2721 bp in length
* 38928: gap of unknown length
* 41649: contig of 3035 bp in length
* 41749: gap of unknown length
* 44783: contig of 2006 bp in length
* 44883: gap of unknown length
* 46889: contig of 2319 bp in length
* 46900: gap of unknown length
* 49308: contig of 2046 bp in length
* 51454: gap of unknown length
* 51455: contig of 3519 bp in length
* 51555: gap of unknown length
* 55074: contig of 3141 bp in length
* 55174: gap of unknown length
* 58315: contig of 3617 bp in length
* 58415: gap of unknown length
* 62032: contig of 4530 bp in length.
* 62132: gap of unknown length
* 62132: contig of 4530 bp in length.
*
* Location/Qualifiers
* 1..66661
* /organism="Bos taurus"
* /mol_type="genomic DNA"
* /db_xref="taxon:9913"
* /clone="CH240-173K8"
* 3788..3887

FEATURES
source
gap

/estimated_length=unknown
6175..6274
/estimated_length=unknown
10189..10288
/estimated_length=unknown
12369..12468
/estimated_length=unknown
17281..17380
/estimated_length=unknown
20214..20313
/estimated_length=unknown
25627..25726
/estimated_length=unknown
27875..27974
/estimated_length=unknown
30255..30354
/estimated_length=unknown
33292..33391
/estimated_length=unknown
38828..38927
/estimated_length=unknown
41649..41748
/estimated_length=unknown
44784..44883
/estimated_length=unknown
46890..46989
/estimated_length=unknown
49309..49408
/estimated_length=unknown
51455..51554
/estimated_length=unknown
55074..55173
/estimated_length=unknown
58315..58414
/estimated_length=unknown
62032..62131
/estimated_length=unknown

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 66661;

Best Local Similarity 94.1%; Pred. No. 2.1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAG 17

Db 29600 GCCACGCTTGGCCGAG 29616

RESULT 57
AY663411
LOCUS
DEFINITION
Homo sapiens voucher Coriell Cell Repository DNA sample NA14663 MHC
class II antigen (HLA-DQB1) gene, HLA-DQB1*0602 allele, MHC
class II antigen (HLA-DQA1) gene, HLA-DQA1*010201 allele, and
MHC class II antigen (HLA-DRB1) gene, complete cds.

ACCESSION AY663411

VERSION AY663411.1 GI:52840236

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo

1 (bases 1 to 99966)

REFERENCE
AUTHORS
Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.

TITLE Ancient haplotypes of the HLA Class II region

JOURNAL Genome Res. 15 (9), 1250-1257 (2005)

PUBMED 16140993

REFERENCE 2 (bases 1 to 99966)

AUTHORS
Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.

TITLE Direct Submission

JOURNAL FEATURES
source

Submitted (18-JUN-2004) Genome Center, Department of Medicine,
University of Washington, Box 352145, Seattle, WA 98195, USA
Location/Qualifiers
1..99966

/organism="Homo sapiens"
/mol_type="genomic DNA"
/specimen_voucher="Coriell Cell Repository DNA sample
NA14663"

/db_xref="taxon:9606"

<3025..>9639

/gene="HLA-DQB1"

/allele="DQB1*0602"

join(<3063..3171,4603..4872,7422..7703,8213..8323,

9421..>9434)

/gene="HLA-DQB1"

/allele="DQB1*0602"

/product="MHC class II antigen"

join(3063..3171,4603..4872,7422..7703,8213..8323,

9421..9434)

/gene="HLA-DQB1"

/allele="DQB1*0602"

/codon_start=1

/product="MHC class II antigen"

/protein_id="AAU88024.1"

/db_xref="GI:52840239"

/translation="MSWKALRIKGLRIVATVTLMLSSLAEGRDSDPEDFVFOFK
CMCYFTNGTERVRLVTRYIYNREYARFSDGVYRAVTPGPRPAEYWNQKEVLEG
TRALDTVCRHNYEVAFRGILQRRVEPTVITSPSTEALNHNHLVCSVTDYDPGQIK
VWFRNDQETAGVSTPLIRNGDWFQILVNLMTPOQGDVYTCHEVPSIQSPITY
EWRAQSEASQSKMLSGVGGFVLGLIFLGLGLIIRQSQKGLLH"
complement(26077..32251)

/gene="HLA-DQA1"

/allele="DQA1*010201"

complement(join(<26903..27057,27414..27695,28109..28357,

32128..>32209))

/gene="HLA-DQA1"

/allele="DQA1*010201"

/product="MHC class II antigen"

complement(join(26903..27057,27414..27695,28109..28357,

32128..>32209))

/gene="HLA-DQA1"

/allele="DQA1*010201"

/codon_start=1

/product="MHC class II antigen"

/protein_id="AAU88022.1"

/db_xref="GI:52840237"

/translation="MILNKALLGALATTVNSPCGGBDVIADHVASCNVLYQYFGP
SQYTHPEFDGEQFYVDLERKETAWRWEFSKFGDFQGALRNMAVAKHNINIKR
YNSTAATNEVPEVTVFSKSPVTLGQPNLILCLVDNIFPPVNIITWLSNQSVTEGVSE
TSFLSKSDHSFFKISYLTFLPSADEIYDKVEHMGLDQPLLKHWEPEIPAPNSELTE
VVCALGLSVGLMGIVGVTFIQLGSRVSGASRHQGPL"
<79861..>90906

/gene="HLA-DRB1"

join(<79937..80036,85301..85570,87837..88118,88819..88929,

89405..89428,90571..>90584)

/gene="HLA-DRB1"

/product="MHC class II antigen"

join(79937..80036,85301..85570,87837..88118,88819..88929,

89405..89428,90571..90584)

/gene="HLA-DRB1"

/codon_start=1

/product="MHC class II antigen"

/protein_id="AAU88023.1"

/db_xref="GI:52840238"

/translation="MVCUKLPGSCMTATVTLMLVSSPLASGSTRFRFLWQPKREK
HFFNGTERVRLDRHFNQESVRFDSVGEFRVATLGRDAEYWNQKQLEQARA
AVDTYCRHNYGVESFTVQRVQPKVTPYPSKTPQLQHNHLVCSVGFYSGSIEVRW
FLNGEEKAGMVTGLIQNGDWTFTLVLETPESGCVYTTCOVHPSPVTSPLTVEWR
ARSEASQSKMLSGVGGFVLGLFLGAGLFIIFRNQKHSGHQPTGFLS"

ORIGIN

Query Match 90.6%; Score 15.4; DB 5; Length 99966;

Best Local Similarity 94.1%; Pred. No. 2.2e+03;

Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

Location/Qualifiers

source

1..103216

/organism="Homo sapiens"

/mol_type="Genomic DNA"

/db_xref="RZPD:RPCIP704P1893"

/db_xref="taxon:9606"

/chromosome="1"

/map="p35.2-36.23"

/clone="RP1-93P18"

/clone_lib="RPCI-1"

misc_feature

1 /note="Clone left end: RP1-93P18"

gene
join(Complement(36883..36946),complement(35500..35595),
complement(34184..34324),complement(32725..32864),
complement(27785..27905),complement(27206..27298),
complement(26164..26294),complement(24361..24488),
complement(5834..6015),complement(5184..5276),
complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:15028..15142),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

/gene="ALDH4A1"

locus_tag="RP11-128M10.1-005"
join(Complement(36883..36946),complement(35500..35595),
complement(34184..34324),complement(32725..32864),
complement(27785..27905),complement(27206..27298),
complement(26164..26294),complement(24361..24488),
complement(5834..6015),complement(5184..5276),
complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:15028..15142),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

/gene="ALDH4A1"

locus_tag="RP11-128M10.1-005"
join(Complement(36883..36946),complement(35500..35595),
complement(34184..34324),complement(32725..32864),
complement(27785..27905),complement(27206..27298),
complement(26164..26294),complement(24361..24488),
complement(5834..6015),complement(5184..5276),
complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:15028..15142),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

/gene="ALDH4A1"

locus_tag="RP11-128M10.1-005"

/product="aldehyde dehydrogenase 4 family, member A1"

/note="match: CDNAs: Em:AY325227.1"

join(Complement(18284..18603),complement(5834..5927),
complement(5184..5276),complement(2286..2333),

complement(AL954340.4:15217..15372),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

/gene="ALDH4A1"

locus_tag="RP11-128M10.1-001"

join(Complement(18284..18603),complement(5834..5927),
complement(5184..5276),complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

/gene="ALDH4A1"

locus_tag="RP11-128M10.1-002"

join(Complement(18284..18375),complement(5834..5927),
complement(5184..5276),complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

/gene="ALDH4A1"

locus_tag="RP11-128M10.1-002"

join(Complement(18284..18375),complement(5834..5927),
complement(5184..5276),complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(BX537160.6:4825..5049),
complement(BX537160.6:3640..3812),
complement(BX537160.6:2681..2968),
complement(AL831755.9:89404..90177))

```
complement(AL954340.4:3548..2701),
complement(BX537160.6:16782..17140))
/gene="ALDH4A1"
/locus tag="RPL1-128M10.1-002"
/product="aldehyde dehydrogenase 4 family, member A1"
/notes="match: cDNAs: Em:U24267.1"
join(complement(18284..18345),complement(5834..5927),
complement(5184..5276),complement(2286..2333),
complement(AL954340.4:15217..15372),
complement(AL954340.4:13023..13172),
complement(AL954340.4:12868..12942),
complement(AL954340.4:11444..11631),
complement(AL954340.4:9044..9117),
complement(AL954340.4:7160..7356),
complement(AL954340.4:6952..6999),
complement(AL954340.4:6059..6211),
complement(AL954340.4:5126..5247),
complement(AL954340.4:4207..4325),
complement(AL954340.4:2589..2701))

Query Match      90.6%; Score 15.4; DB 5; Length 103216;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCACGCTTGCCGCGAG 17
|||||
Db 43206 GCCCACGCTTGCCCGAG 43190

RESULT 60
AY663406
LOCUS
DEFINITION
AY663406 Homo sapiens voucher Coriell Cell Repository DNA sample NA10923 MHC
class II antigen (HLA-DQB1) gene, HLA-DQB1*0602 allele, MHC
class II antigen (HLA-DQA1) gene, HLA-DQA1*010201 allele, and
MHC class II antigen (HLA-DRB1) gene, HLA-DRB1*150101 allele,
complete cds.
ACCESSION
AY663406
VERSION
AY663406.1 GI:52840216
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 103395)
Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
2 (bases 1 to 103395)
Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.
Ancient haplotypes of the HLA Class II region
Genome Res. 15 (9), 1250-1257 (2005)
16140993
3 (bases 1 to 103395)
Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.
Direct Submission
Submitted (18-JUN-2004) Genome Center, Department of Medicine,
University of Washington, Box 352145, Seattle, WA 98195, USA
Location/Qualifiers
1..103395
/organism="Homo sapiens"
/mol_type="genomic DNA"
/specimen_voucher="Coriell Cell Repository DNA sample
NA10923"
/db_xref="taxon:9606"
<13398..>20612
/gene="HLA-DQB1"
/allele="DQB1*0602"
join(<14036..14144,15576..15845,18395..18676,19186..19296,
20394..>20407)
/gene="HLA-DQB1"
/allele="DQB1*0602"
/product="MHC class II antigen"
join(14036..14144,15576..15845,18395..18676,19186..19296,
```

```
20394..20407)
/gene="HLA-DQB1"
/allele="DQB1*0602"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAU88009.1"
/db_xref="GI:52840219"
/translation="MSWKALRIIPGDLRVATVTLMLAMLSLLAEGRDSPEDFVQFQK
GMCYFNGTGERVLTVIYNYREEYARFSDVGVYRAVTPQGRDAIYWNYSQKVLGK
TRAEIDTCRHNYEYVAFRGLIQORVEPTVITSPSRTEALHNHLLVGSVTDIFYGQIK
VVRFRNDQEBETAGVSTPLIRNGDWTFOILLMLEMTFQRGDVYTCHVEHPSLQSPITV
EWRAQESAKMSLGVGGFVLGLIFLGLLIIRQSQKGLLH"
complement(37053..43226)
/gene="HLA-DQA1"
/allele="DQA1*010201"
complement(join(<37879..38033,38390..38671,39085..39333,
43103..>43184))
/gene="HLA-DQA1"
/allele="DQA1*010201"
/product="MHC class II antigen"
complement(join(37879..38033,38390..38671,39085..39333,
43103..43184))
/gene="HLA-DQA1"
/allele="DQA1*010201"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAU88007.1"
/db_xref="GI:52840217"
/translation="MILNKALLIGALALTTVMSPGGEDIIVADHVASCNVLYQFYGP
SGOYTFEFGDEQFYDLERKETAWRPFSKFGGFPQOGLRMAVAKHNLNIMIKR
YNSATANEVPEVTFSKSVTLGQPNLTLCLVDNI PFPVNIITLNGQSQSVTEGVSE
TSFLSKSDHSDFFKISYLTFLPSADEIYDKVEHWGLDQPLLKHWEPIAPMSELTET
VVCALGSLVMGIIVGVTFIIQGLRSVGASRHQGPL"
<90424..>101471
/gene="HLA-DRB1"
/allele="DRB1*150101"
join(<90500..90599,95864..96133,98402..98683,99384..99494,
99970..99993,101136..>101149)
/gene="HLA-DRB1"
/allele="DRB1*150101"
/product="MHC class II antigen"
join(90500..90599,95864..96133,98402..98683,99384..99494,
99970..99993,101136..101149)
/gene="HLA-DRB1"
/allele="DRB1*150101"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAU88008.1"
/db_xref="GI:52840218"
/translation="MVCLKLPGSCMTALTITLMLVSLSPALSGLDTRPRFLWQPKREC
HFNGTERTVPLDRIFYNQBESEVRFSDVGEFRAVTELGRLDAIYWNYSQKDIIEQARA
AVDTYCRHNYGVVESFTVQRVQPKVTVPYPSKTPQLQHNLNLLVGSVGFYFGSIEVRW
FLNGQESKAGMVGTLIQNGDWTFTLMLVETVPRSEVYTCQVEHPSVTPLTVEWR
ARSESAQSKMSLGVGGFVLGLIFLGLLIIFVRNQKHSGLQPTGFLS"

ORIGIN
Query Match      90.6%; Score 15.4; DB 5; Length 103395;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCACGCTTGCCGCGAG 17
|||||
Db 15871 GCCCACCTTGCCCGAG 15887

RESULT 61
AY663395
LOCUS
DEFINITION
AY663395 Homo sapiens voucher Coriell Cell Repository DNA sample NA14660 MHC
class II antigen (HLA-DQB1) gene, HLA-DQB1*0602 allele, MHC
class II antigen (HLA-DQA1) gene, HLA-DQA1*010201 allele, and
MHC class II antigen (HLA-DRB1) gene, complete cds.
ACCESSION
AY663395
```

VERSION AY663395.1 GI:52840176
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 104996)
AUTHORS Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.
TITLE Ancient haplotypes of the HLA Class II region
JOURNAL Genome Res. 15 (9), 1250-1257 (2005)
PUBMED 16140993
REFERENCE 2 (bases 1 to 104996)
AUTHORS Raymond,C.K., Kas,A., Paddock,M., Qiu,R., Zhou,Y., Subramanian,S.,
Chang,J., Palmieri,A., Haugen,E., Kaul,R. and Olson,M.V.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2004) Genome Center, Department of Medicine,
University of Washington, Box 352145, Seattle, WA 98195, USA
FEATURES
source
1..104996
/organism="Homo sapiens"
/mol_type="genomic DNA"
/specimen_voucher="Coriell Cell Repository DNA sample
NA14660"
/db_xref="taxon:9606"
<2884..>9498
/gene="HLA-DQB1"
/alleles="DQB1*0602"
join(<2922..3030,4462..4731,7281..7562,8072..8182,
9280..>9293)
/gene="HLA-DQB1"
/alleles="DQB1*0602"
/product="MHC class II antigen"
join(2922..3030,4462..4731,7281..7562,8072..8182,
9280..>9293)
/gene="HLA-DQB1"
/alleles="DQB1*0602"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAU87980.1"
/db_xref="GI:52840179"
translation="MSWKRLRIPGDLRVATVTLMLSLAERDSDPFDVFPQK
GMVFTNCTERVLRLVYIYNREVARSDVGVRVTPQGRPDAYWNSQKVLG
TRALDTVCRNHYEVAFGILQRVEPTVTSRTEALHNHLLVCSVTDFYGGIK
WFRNDOEETAGVSTPLIRNGDWTFQTLVLMLETPQGRDVTYCHVEHPSLQSPITV
EWRAQSAQSKMLSGVGGFVLGLIFLGLGLIIRQSRQKGLLH"
complement(25936..32110)
/gene="HLA-DQA1"
/alleles="DQA1*010201"
complement(join(<26762..26916,27273..27554,27968..28216,
31987..>32068))
/gene="HLA-DQA1"
/alleles="DQA1*010201"
/product="MHC class II antigen"
complement(join(26762..26916,27273..27554,27968..28216,
31987..>32068))
/gene="HLA-DQA1"
/alleles="DQA1*010201"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAU87978.1"
/db_xref="GI:52840177"
translation="MLNKALLGALATVWSPGGEDIVADHVASGVNLYFYFGP
SQQYTFDEQEYDLERKETARWPEFSKFGDFDQGARNNVAKHNLNIMIKR
YNSAATNEDVEPVFVSKSPVTLGQNTLCLVDNIPFPVNVITWLSNGQSVTEGVSE
TSFLSKDSHFFKISYLTFLPSADEIYDKVEHMGLOPLLKWPEIPAPMSELLET
VVCALGSLVGLMGLVVGTVFIQLGRSVGASRHQPL"
<79388..>90431
/gene="HLA-DRB1"
join(<79464..79563,84828..85097,87362..87643,88344..88454,
88930..88953,90096..>90109)
/gene="HLA-DRB1"

CDS
/product="MHC class II antigen"
join(79464..79563,84828..85097,87362..87643,88344..88454,
88930..88953,90096..90109)
/gene="HLA-DRB1"
/codon_start=1
/product="MHC class II antigen"
/protein_id="AAU87979.1"
/db_xref="GI:52840178"
translation="MVCLKPGGSCMTALTVTLMVLSPLASGDTRRPFLWQPKREC
HFFNGTERVFLDRHFYNQESVRPDSVGEFRAVTELGPRDAEYWNQSKDLQEQARA
AVDTYCRHNYGVESFTVRRVQPKVTVPKTPQLQHNLLVCSVGFYPGSIEVRW
FLNQEEKAGMVGSTGLIQNGDWTFQTLVLMLETPRSGEVYTCQVHPSVTSPLTVEWR
ARSESAQSKMLSGVGGFVLGLFLGAGLFIYFRNQKHGSLQPTGFLS"
ORIGIN
Query Match 90.6%; Score 15.4; DB 5; Length 104996;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCACGCTTGGCCGAG 17
Db 4757 GCCCACCTTGGCCGAG 4773
RESULT 62
AE001439_02/c
WPCOMMENT
Sequence split into 17 fragments LOCUS AE001439 Accession AE001439
Fragment Name Begin End
AE001439_00 1 110000
AE001439_01 100001 210000
AE001439_02 200001 310000
AE001439_03 300001 410000
AE001439_04 400001 510000
AE001439_05 500001 610000
AE001439_06 600001 710000
AE001439_07 700001 810000
AE001439_08 800001 910000
AE001439_09 900001 1010000
AE001439_10 1000001 1110000
AE001439_11 1100001 1210000
AE001439_12 1200001 1310000
AE001439_13 1300001 1410000
AE001439_14 1400001 1510000
AE001439_15 1500001 1610000
AE001439_16 1600001 1643831
Continuation (3 of 17) of AE001439 from base 200001 (AE001439 Helicobacter pylori J99, c
Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCACGCTTGGCCGAG 17
Db 109190 GCCCACGCTTGGCCGAG 109174
RESULT 63
AE001439_03/c
WPCOMMENT
Sequence split into 17 fragments LOCUS AE001439 Accession AE001439
Fragment Name Begin End
AE001439_00 1 110000
AE001439_01 100001 210000
AE001439_02 200001 310000
AE001439_03 300001 410000
AE001439_04 400001 510000
AE001439_05 500001 610000
AE001439_06 600001 710000
AE001439_07 700001 810000
AE001439_08 800001 910000
AE001439_09 900001 1010000
AE001439_10 1000001 1110000

AE001439_11 1100001 1210000
AE001439_12 1200001 1310000
AE001439_13 1300001 1410000
AE001439_14 1400001 1510000
AE001439_15 1500001 1610000
AE001439_16 1600001 1643831
Continuation (4 of 17) of AE001439 from base 300001 (AE001439 Helicobacter pylori J99, c

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
| | | | | | | | | | | | | | | | | | | |
Db 9190 GCCCAGCGTTGGCCGAG 9174

RESULT 64
CP000240 27/c
WPCOMMENT
Sequence split into 31 fragments LOCUS CP000240 Accession CP000240

Fragment Name	Begin	End
CP000240_00	1	110000
CP000240_01	100001	210000
CP000240_02	200001	310000
CP000240_03	300001	410000
CP000240_04	400001	510000
CP000240_05	500001	610000
CP000240_06	600001	710000
CP000240_07	700001	810000
CP000240_08	800001	910000
CP000240_09	900001	1010000
CP000240_10	1000001	1110000
CP000240_11	1100001	1210000
CP000240_12	1200001	1310000
CP000240_13	1300001	1410000
CP000240_14	1400001	1510000
CP000240_15	1500001	1610000
CP000240_16	1600001	1710000
CP000240_17	1700001	1810000
CP000240_18	1800001	1910000
CP000240_19	1900001	2010000
CP000240_20	2000001	2110000
CP000240_21	2100001	2210000
CP000240_22	2200001	2310000
CP000240_23	2300001	2410000
CP000240_24	2400001	2510000
CP000240_25	2500001	2610000
CP000240_26	2600001	2710000
CP000240_27	2700001	2810000
CP000240_28	2800001	2910000
CP000240_29	2900001	3010000
CP000240_30	3000001	3046682

Continuation (29 of 31) of CP000240 from base 2800001 (CP000240 Cyanobacteria bacterium

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
| | | | | | | | | | | | | | | | | | | |
Db 616 GCCCAGCGTTGGCCGAG 600

RESULT 66
CP000248 28/c
WPCOMMENT
Sequence split into 36 fragments LOCUS CP000248 Accession CP000248

Fragment Name	Begin	End
CP000248_00	1	110000
CP000248_01	100001	210000
CP000248_02	200001	310000
CP000248_03	300001	410000
CP000248_04	400001	510000
CP000248_05	500001	610000
CP000248_06	600001	710000
CP000248_07	700001	810000
CP000248_08	800001	910000
CP000248_09	900001	1010000
CP000248_10	1000001	1110000
CP000248_11	1100001	1210000
CP000248_12	1200001	1310000
CP000248_13	1300001	1410000
CP000248_14	1400001	1510000
CP000248_15	1500001	1610000
CP000248_16	1600001	1710000
CP000248_17	1700001	1810000
CP000248_18	1800001	1910000
CP000248_19	1900001	2010000
CP000248_20	2000001	2110000
CP000248_21	2100001	2210000
CP000248_22	2200001	2310000
CP000248_23	2300001	2410000
CP000248_24	2400001	2510000
CP000248_25	2500001	2610000
CP000248_26	2600001	2710000
CP000248_27	2700001	2810000
CP000248_28	2800001	2910000
CP000248_29	2900001	3010000

Continuation (28 of 31) of CP000240 from base 2700001 (CP000240 Cyanobacteria bacterium

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
| | | | | | | | | | | | | | | | | | | |
Db 100616 GCCCAGCGTTGGCCGAG 100600

RESULT 65
CP000240 28/c
WPCOMMENT
Sequence split into 31 fragments LOCUS CP000240 Accession CP000240

Fragment Name	Begin	End
CP000240_00	1	110000
CP000240_01	100001	210000
CP000240_02	200001	310000
CP000240_03	300001	410000

CP000248_30 3000001 3110000
CP000248_31 3100001 3210000
CP000248_32 3200001 3310000
CP000248_33 3300001 3410000
CP000248_34 3400001 3510000
CP000248_35 3500001 3561584
Continuation (29 of 36) of CP000248 from base 2800001 (CP000248 Novosphingobium aromaticum)

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17

Db 72861 GCCCAGCTTGGCCGAG 72845

RESULT 67

AE014295_08/c

WPCOMMENT

Sequence split into 23 fragments LOCUS AE014295 Accession AE014295

Fragment Name Begin End

AE014295_00 1 110000
AE014295_01 100001 210000
AE014295_02 200001 310000
AE014295_03 300001 410000
AE014295_04 400001 510000
AE014295_05 500001 610000
AE014295_06 600001 710000
AE014295_07 700001 810000
AE014295_08 800001 910000
AE014295_09 900001 1010000
AE014295_10 1000001 1110000
AE014295_11 1100001 1210000
AE014295_12 1200001 1310000
AE014295_13 1300001 1410000
AE014295_14 1400001 1510000
AE014295_15 1500001 1610000
AE014295_16 1600001 1710000
AE014295_17 1700001 1810000
AE014295_18 1800001 1910000
AE014295_19 1900001 2010000
AE014295_20 2000001 2110000
AE014295_21 2100001 2210000
AE014295_22 2200001 2256640
Continuation (9 of 23) of AE014295 from base 800001 (AE014295 Bifidobacterium longum NC0

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17

Db 24523 GCCCAGCTTGGCCGAG 24507

RESULT 68

AE017262_13/c

WPCOMMENT

Sequence split into 29 fragments LOCUS AE017262 Accession AE017262

Fragment Name Begin End

AE017262_00 1 110000
AE017262_01 100001 210000
AE017262_02 200001 310000
AE017262_03 300001 410000
AE017262_04 400001 510000
AE017262_05 500001 610000
AE017262_06 600001 710000
AE017262_07 700001 810000
AE017262_08 800001 910000
AE017262_09 900001 1010000
AE017262_10 1000001 1110000
AE017262_11 1100001 1210000

AE017262_12 1200001 1310000
AE017262_13 1300001 1410000
AE017262_14 1400001 1510000
AE017262_15 1500001 1610000
AE017262_16 1600001 1710000
AE017262_17 1700001 1810000
AE017262_18 1800001 1910000
AE017262_19 1900001 2010000
AE017262_20 2000001 2110000
AE017262_21 2100001 2210000
AE017262_22 2200001 2310000
AE017262_23 2300001 2410000
AE017262_24 2400001 2510000
AE017262_25 2500001 2610000
AE017262_26 2600001 2710000
AE017262_27 2700001 2810000
AE017262_28 2800001 2905187

Continuation (14 of 29) of AE017262 from base 1300001 (AE017262 Listeria monocytogenes

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17

Db 31052 GCCCAGCTTGGCTGAG 31036

RESULT 69

BA000032_12/c

WPCOMMENT

Sequence split into 19 fragments LOCUS BA000032 Accession BA000032

Fragment Name Begin End

BA000032_00 1 110000
BA000032_01 100001 210000
BA000032_02 200001 310000
BA000032_03 300001 410000
BA000032_04 400001 510000
BA000032_05 500001 610000
BA000032_06 600001 710000
BA000032_07 700001 810000
BA000032_08 800001 910000
BA000032_09 900001 1010000
BA000032_10 1000001 1110000
BA000032_11 1100001 1210000
BA000032_12 1200001 1310000
BA000032_13 1300001 1410000
BA000032_14 1400001 1510000
BA000032_15 1500001 1610000
BA000032_16 1600001 1710000
BA000032_17 1700001 1810000
BA000032_18 1800001 1877212
Continuation (13 of 19) of BA000032 from base 1200001 (BA000032 Vibrio parahaemolyticus

Query Match 90.6%; Score 15.4; DB 15; Length 110000;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17

Db 33479 GCCCAGCTTGGCCAAG 33463

RESULT 70

AP003159/c

LOCUS ..

DEFINITION Homo sapiens genomic DNA, chromosome 1p36 clone:1024F16, complete

sequence.

ACCESSION AP003159

VERSION AP003159.1

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1
Machida,T., Ohira,M., Morohashi,A., Mihara,M., Furuta,S., Soeda,E. and Nakagawara,A.
Homo sapiens 113,733 genomic DNA of 1p36
Published Only in Database (2001)

2 (bases 1 to 113733)
Machida,T., Ohira,M., Morohashi,A., Mihara,M., Furuta,S., Soeda,E. and Nakagawara,A.
Direct Submission
Submitted (31-JAN-2001) Akira Nakagawara, Chiba Cancer Center Research Institute, Division of Biochemistry; 666-2 Nitona, Chuoh-ku, Chiba, Chiba 260-8717, Japan
(E-mail: akiranak@chiba-cc.pref.chiba.jp, Tel: 81-43-264-5431, Fax: 81-43-265-4459)
Additional author information
Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
The Institute of Physical and Chemical Research (RIKEN) Genomic Sciences Center (GSC)
1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa, JAPAN zip: 230-0045
phone: 81-45-503-9111, fax: 81-45-503-9170
e-mail: hattori@gs.c.riken.go.jp.
Location/Qualifiers
1. 113733
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36"
/clone="1024F16"

FEATURES
source
1. 147557
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clones="XXbac-254F23"
/clone_lib="CHORI-501"

misc_feature
1
/note="Clone_left_end: XXbac-254F23"
6890..7838
/note="Single clone region. Sequence from reads from a short insert library and a transposon library derived from a single pUC clone. Restriction digest data confirm the assembly."
44278
/note="Clone_right_end: XXbac-161M6"
49251..55484
/gene="HLA-DQA1"
/locus_tag="XXbac-BPG254F23.1-001"
join(49251..49434,53204..53452,53866..54147,54504..54678,54846..55484)
/gene="HLA-DQA1"

gene
1
/locus_tag="XXbac-BPG254F23.1-001"
/note="match: ESTs: AA381553.1 AV717599.1 BQ049081.1 match: CDNAS: AF037314.1 BC008585.1 M29939.1 M33906.1" join(49353..49434,53204..53452,53866..54147,54504..54658)
/gene="HLA-DQA1"

mRNA
1
/locus_tag="XXbac-BPG254F23.1-001"
/standard_name="OTHUMP00000029141"
/note="match: proteins: O19760 P01909 P04226 Q30410 Q30438"
/codon_start=1
/protein_id="CAI18230.1"
/db_xref="GI:55961973"
/db_xref="GOA:Q5V7D5"
/db_xref="InterPro:IPR001003"
/db_xref="InterPro:IPR003006"
/db_xref="InterPro:IPR003597"
/db_xref="UniProtKB/TrEMBL:Q5V7D5"
/translation="MILNKALLGALATTVMSPCGGEDIVADHVASCGLNYQFYGP SGQYTFDGDGEQFYVDLERKETAWRPFEKFGFQDQALRNWAVAKHNLINIKR YNSTAATNEVEPFFVFSKSPVTLGPNTLICLVNDIPFPVNVITWLSNGQSVTEGVSE TFSLSKSDHSFKISYILTFPSADEIYDCKVEHWGLDQDLKHWEPFAPPMSELTET VVCALGLSVGLMGIVVGTFTFIQGLRSVGSRRHQPL"

CDS
1
/locus_tag="XXbac-BPG254F23.1-001"
/standard_name="OTHUMP00000029141"
/note="match: proteins: O19760 P01909 P04226 Q30410 Q30438"
/codon_start=1
/protein_id="CAI18230.1"
/db_xref="GI:55961973"
/db_xref="GOA:Q5V7D5"
/db_xref="InterPro:IPR001003"
/db_xref="InterPro:IPR003006"
/db_xref="InterPro:IPR003597"
/db_xref="UniProtKB/TrEMBL:Q5V7D5"
/translation="MILNKALLGALATTVMSPCGGEDIVADHVASCGLNYQFYGP SGQYTFDGDGEQFYVDLERKETAWRPFEKFGFQDQALRNWAVAKHNLINIKR YNSTAATNEVEPFFVFSKSPVTLGPNTLICLVNDIPFPVNVITWLSNGQSVTEGVSE TFSLSKSDHSFKISYILTFPSADEIYDCKVEHWGLDQDLKHWEPFAPPMSELTET VVCALGLSVGLMGIVVGTFTFIQGLRSVGSRRHQPL"

REFERENCE
1
Machida,T., Ohira,M., Morohashi,A., Mihara,M., Furuta,S., Soeda,E. and Nakagawara,A.
Homo sapiens 113,733 genomic DNA of 1p36
Published Only in Database (2001)

2 (bases 1 to 113733)
Machida,T., Ohira,M., Morohashi,A., Mihara,M., Furuta,S., Soeda,E. and Nakagawara,A.
Direct Submission
Submitted (31-JAN-2001) Akira Nakagawara, Chiba Cancer Center Research Institute, Division of Biochemistry; 666-2 Nitona, Chuoh-ku, Chiba, Chiba 260-8717, Japan
(E-mail: akiranak@chiba-cc.pref.chiba.jp, Tel: 81-43-264-5431, Fax: 81-43-265-4459)
Additional author information
Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
The Institute of Physical and Chemical Research (RIKEN) Genomic Sciences Center (GSC)
1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa, JAPAN zip: 230-0045
phone: 81-45-503-9111, fax: 81-45-503-9170
e-mail: hattori@gs.c.riken.go.jp.
Location/Qualifiers
1. 113733
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36"
/clone="1024F16"

ORIGIN
Query Match 90.6%; Score 15.4; DB 5; Length 113733;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACGCTTGGCCGAG 17
|||||
Db 44326 GCCACGCTTGGCCGAG 44310
|||||

RESULT 71
AL662789/c
LOCUS
DEFINITION
Human DNA sequence from clone XXbac-254F23 on chromosome 6 contains the 5' end of the HLA-DRB1 gene for major histocompatibility complex, class II, DR beta 1, the HLA-DQA1 gene for major histocompatibility complex, class II, DQ alpha 1, the HLA-DQB1 gene for major histocompatibility complex, class II, DQ beta 1, a cytochrome C oxidase polypeptide III pseudogene, a major histocompatibility complex, class II, beta polypeptide pseudogene and two CpG islands, complete sequence.
AL662789
GI:19808179
HTG; CpG island; cytochrome C oxidase polypeptide III; DQA1; HLA-DQ; HLA-DQA; HLA-DQA1; HLA-DQB; HLA-DQB1; HLA-DRB1; HLA-DRB1; IDDM1; major histocompatibility complex, class II, beta; major histocompatibility complex, class II, DQ alpha; major histocompatibility complex, class II, DQ beta; major histocompatibility complex, class II, DR beta.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 147557)
Tracey,A.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

COMMENT
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
On Mar 29, 2002 this sequence version replaced gi:19808178.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
XXbac-254F23 is from a CHORI-501 human bac - PGF cell line library VECTOR: pTARBAC2.1
This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1. 147557
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clones="XXbac-254F23"
/clone_lib="CHORI-501"

```
gene      complement(71364..78711)
          /gene="HLA-DQB1"
          /locus tag="XXbac-BPG254F23.2-001"
mRNA      complement(join(71364..72143,73241..73351,73861..74142,
76692..76961,78393..78711))
          /gene="HLA-DQB1"
          /locus tag="XXbac-BPG254F23.2-001"
          /note="match: ESTs: BG542046.1"
          match: cDNAs: M16276.1 M17955.1 M25325.1 M25327.1 M32577.1
X03068.1"
CDS       complement(join(72130..72143,73241..73351,73861..74142,
76692..76961,78393..78501))
          /gene="HLA-DQB1"
          /locus tag="XXbac-BPG254F23.2-001"
          /standard name="ORTHUMP0000029167"
          /note="match: proteins: AAH12106 P01919 Q29884 Q29967
Q30061 Q30091 Q30096"
          /codon_start=1
          /protein_id="CAI18231.1"
          /db_xref="GI:55961974"
          /db_xref="InterPro:IPR000353"
          /db_xref="InterPro:IPR003006"
          /db_xref="InterPro:IPR003597"
          /db_xref="InterPro:IPR007110"
          /translation="MSMKALRIIPGDLRVATVTLMLSLLAEGRDSPEDFQFQFK
GMCYFTNGTERVRLVRIYINREVEAFDSGVVRAVTPQGRPAEYVNSQKEVLEG
TRAEIDTCRNIYEVARRGILQRRVEPTVTSRTEALNHNLLVCSTVDFPQGIK
VWFRNDEETAGVSTPLIRNGDWTFCILVPLENTFQGRDVTYCHVEHPSLQSPITV
EWRAQSAQKMLSGVGVGLIFGLIGLIIRQSRQKGLLH"
          /complement(118016..118697)
          /locus tag="XXbac-BPG254F23.4-001"
          /pseudo
CDS       complement(118016..118697)
          /locus tag="XXbac-BPG254F23.4-001"
          /note="match: proteins: O99810 Q35486 Q36378 Q92J29 Q94QC1
Q94TJ9 Q94285 Q9G3E4 Q9G3R7"
          /pseudo
          /codon_start=1
          /product="cytochrome c oxidase III pseudogene 1"
          complement(143223..143590)
          /locus tag="XXbac-BPG254F23.3-001"
          /pseudo
          complement(143223..143590)
          /locus tag="XXbac-BPG254F23.3-001"
          /note="match: proteins: O19507 O62868 O98263 Q85PC4 Q95399
Q98S54"
          /pseudo
          /codon_start=1
          /product="major histocompatibility complex, class II, DQ
beta 3"
          147557
          /note="Clone_right_end: XXbac-254F23"
misc_feature
          147557
ORIGIN
Query Match      90.6%; Score 15.4; DB 5; Length 147557;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 GCCCAGCGTTGCCGAG 17
          ||||| ||||| |||||
Db      76666 GCCCACCCTTGCCGAG 76650

RESULT 72
AC150103/c      AC150103      150372 bp      DNA      linear      HTG 01-JUL-2004
LOCUS      Gallus gallus clone WAG-102C19, WORKING DRAFT SEQUENCE, 18 ordered
DEFINITION      pieces.
ACCESSION      AC150103
VERSION      AC150103.1 GI:49533698
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus
1 (bases 1 to 150372)
DOE Joint Genome Institute.
Unpublished
2 (bases 1 to 150372)
DOE Joint Genome Institute.
Direct Submission
Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
94598-1698, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 2390506
Center clone name: Chk_102C19

Summary Statistics
Consensus quality: 141126 bases at least Q40
Consensus quality: 145414 bases at least Q30
Consensus quality: 147452 bases at least Q20
Estimated insert size: 171000; agarose-fp estimation
Estimated insert size: 148672; sum-of-contigs estimation
Quality coverage: 11.15 in Q20 bases; agarose-fp estimation
Quality coverage: 12.83 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 34543: contig of 34543 bp in length
* 34544 34643: gap of unknown length
* 34644 38244: contig of 3601 bp in length
* 38245 38344: gap of unknown length
* 38345 44580: contig of 6236 bp in length
* 44581 44680: gap of unknown length
* 44681 45294: contig of 4614 bp in length
* 45295 49394: gap of unknown length
* 49395 59164: contig of 9770 bp in length
* 59165 59264: gap of unknown length
* 59265 65805: contig of 6541 bp in length
* 65806 65905: gap of unknown length
* 65906 80351: contig of 14446 bp in length
* 80352 80451: gap of unknown length
* 80452 83731: contig of 3280 bp in length
* 83732 83831: gap of unknown length
* 83832 96414: contig of 12583 bp in length
* 96415 96514: gap of unknown length
* 96515 99192: contig of 2878 bp in length
* 99193 100721: contig of 1429 bp in length
* 100722 100821: gap of unknown length
* 100822 102609: contig of 1788 bp in length
* 102610 102709: gap of unknown length
* 102710 110148: contig of 7439 bp in length
* 110149 110248: gap of unknown length
* 110249 116540: contig of 6292 bp in length
* 116541 116640: gap of unknown length
* 116641 135574: contig of 18934 bp in length
* 135575 135674: gap of unknown length
* 135675 136513: contig of 839 bp in length
* 136514 136614: gap of unknown length
* 136614 138347: contig of 1734 bp in length
* 138348 138447: gap of unknown length

```

* 138448 150372: contig of 11925 bp in length.
FEATURES
  source      Location/Qualifiers
1. 150372
   /organism="Gallus gallus"
   /mol_type="genomic DNA"
   /db_xref="taxon:9031"
   /clone_lib="Texas A&M Wageningen Chicken BAC Library"
   /clone_id="34544"
   /estimated_length=unknown
gap
38245..38344
   /estimated_length=unknown
gap
44581..44680
   /estimated_length=unknown
gap
49295..49394
   /estimated_length=unknown
gap
59165..59264
   /estimated_length=unknown
gap
65806..65905
   /estimated_length=unknown
gap
80352..80451
   /estimated_length=unknown
gap
83732..83831
   /estimated_length=unknown
gap
96415..96514
   /estimated_length=unknown
gap
99193..99292
   /estimated_length=unknown
gap
100722..100821
   /estimated_length=unknown
gap
102610..102709
   /estimated_length=unknown
gap
110149..110248
   /estimated_length=unknown
gap
116541..116640
   /estimated_length=unknown
gap
135575..135674
   /estimated_length=unknown
gap
136514..136613
   /estimated_length=unknown
gap
138348..138447
   /estimated_length=unknown
ORIGIN
Query Match      90.6%; Score 15.4; DB 12; Length 150372;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
   |||||
Db 116985 GCCCAGCTTGGCCGAG 116969

RESULT 73
AC169542
LOCUS
DEFINITION Bos taurus clone CH240-242B23, WORKING DRAFT SEQUENCE, 9 unordered
pieces.
ACCESSION AC169542
VERSION AC169542.2 GI:85678605
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 150375)
Munzy,D., Adams,C., Agbai II,O., Allen,C., Albrooks,S., Archer,P.,
Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,
Beraducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
Chacko,J., Chahrouh,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
Dziulida,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haeblerlen,K.,
Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Haviak,P.,
Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
Hines,S., Hitchens,M., Hodgson,A., Hogue,M., Hollins,B.,
Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,
Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,E.,
Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
London,P., Lopez,J., Lorensuhewa,L., Lozado,R., Luk,T., Madu,R.,
Maheehwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
McClelland,H., McPherson,J., Mercadao,C., Metzker,M.,
Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Muidasa,M.,
Murray,D., Nazarith,L., Ngo,D., Nguyen,N., Norwig-Estaugh,E.,
Nott,A., Nwaokemele,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,
Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J.,
Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R.,
Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
Trejos,Z., Usmani,K., Vargo,C., Verdusco,D., Villasana,D., Virk,D.,
Voikov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
Wu,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 150375)
Worley,K.C.
Direct Submission
Submitted (07-OCT-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 150375)
Bovine Genome Sequencing Consortium
Direct Submission
Submitted (24-JAN-2006) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jan 24, 2006 this sequence version replaced gi:77377088.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FNFG
Center clone name: CH240-242B23
----- Summary Statistics
Assembly program: Atlas 3.0;

```

Consensus quality: 147151 bases at least Q40
 Consensus quality: 147915 bases at least Q30
 Consensus quality: 148615 bases at least Q20
 Estimated insert size: 153523; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1995: contig of 1995 bp in length
 1 2045: gap of 50 bp
 2046 33670: contig of 31625 bp in length
 33671 33720: gap of 50 bp
 33721 92544: contig of 58824 bp in length
 92545 92550: gap of 106 bp
 92551 100413: contig of 7763 bp in length
 100414 100463: gap of 50 bp
 100464 103962: contig of 3499 bp in length
 103963 104012: gap of 50 bp
 104013 123918: contig of 19906 bp in length
 123919 123968: gap of 50 bp
 123969 146272: contig of 22304 bp in length
 146273 146777: gap of 505 bp
 146778 148829: contig of 2052 bp in length
 148830 148929: gap of unknown length
 148930 150375: contig of 1446 bp in length.

FEATURES

source
 1..150375
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="CH240-242B23"
 1996..2045
 /estimated_length=50
 33671..33720
 /estimated_length=50
 92545..92550
 /estimated_length=106
 100414..100463
 /estimated_length=50
 103963..104012
 /estimated_length=50
 123919..123968
 /estimated_length=50
 146273..146777
 /estimated_length=505
 148830..148929
 /estimated_length=unknown

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 150375;
 Best Local Similarity 94.1%; Pred. No. 2.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GCCACGCTGGCCGAG 17
 |||||
 Db 136558 GCCACGCTGGCCGAG 136574

RESULT 74
 AL355273

LOCUS AL355273 153108 bp DNA linear PRI 18-MAY-2005
 DEFINITION Human DNA sequence from clone RP11-714M16 on chromosome 10. Contains
 the 3' end of the GPRK5 gene for G protein-coupled receptor kinase
 5 (GRK5) and the RGS10 gene for regulator of G-protein signalling
 10, complete sequence.

ACCESSION

AL355273
 AL355273.30 GI:17973918
 KEYWORDS HTG; GPRK5; GRK5; RGS10.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 153108)
 Kay.M.
 Direct Submission

REFERENCE

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Dec 20, 2001 this sequence version replaced gi:17221178.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em, EMBL; Swi, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
 RP11-714M16 is from the library RPCI-11.3 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone; and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES

source

Location/Qualifiers
 1..153108
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-714M16"
 /clone_lib="RPCI-11.3"

misc_feature

2000
 /note="Clone right_end: RP11-79M19"

misc_feature

14785..15226
 /note="Single clone region. Assembly consistent with
 restriction digest data."

misc_feature

14943
 /note="Tandem repeat. Forced join. Gap size estimated to
 be approximately 500bp by restriction digest data."
 join(ALJ355861.19:48518..48898,ALJ355861.19:167445..167540,
 ALJ583824.16:55217..55329,17397..17474,43868..43968,
 45695..45787,51077..51140,52089..52229,57353..57543,
 60433..60470,62701..62790,64246..64454,68825..68962,
 73373..73510,73847..73978,75671..76321)
 /gene="GPRK5"

gene

/locus_tag="RP11-714M16.1-002"
 join(ALJ355861.19:48518..48898,ALJ355861.19:167445..167540,
 ALJ583824.16:55217..55329,17397..17474,43868..43968,
 45695..45787,51077..51140,52089..52229,57353..57543,
 60433..60470,62701..62790,64246..64454,68825..68962,
 73373..73510,73847..73978,75671..76321)
 /gene="GPRK5"
 /locus_tag="RP11-714M16.1-002"
 /product="G protein-coupled receptor kinase 5"
 /note="match: ESTs: AI160744.1 AI338666.1 AI346265.1

mrna

```

AI672434.1 A1738477.1 A1830135.1 A1132088.1 A12999685.1
BF439363.1 BF569124.1 BF972966.1 BG477593.1 BG764871.1
B123269.1 B1333586.1 B14049553.1 B1562122.1 B1563509.1
BQ029666.1 BQ055527.1 BQ056240.1 BQ065277.1 BQ066445.1
BQ071007.1 BQ184611.1 BQ186585.1 BQ18795.1 BQ927452.1
BQ929071.1 BQ935218.1 BQ938926.1 BQ9527897.1 BQ9681200.1
match: cDNAs: L15386.1"
join(ALJ355861.19:48847. .48898,ALJ355861.19:167445. .167540,
ALJ583824.16:55217. .55329,17397. .17474,43868. .43968,
45695. .45787,51077. .51140,52089. .52229,57353. .57543,
60433. .60470,62701. .62790,64246. .64454,68825. .68962,
73373. .73510,73847. .73978,75671. .75769)
/gene="GPRK5"
/locus tag="RP11-714M16.1-002"
/standard_name="OTTHUMP00000020596"
/note="match: proteins: P34947 P43249 Q62833 Q8VEB1"
/codon_start=1
/product="G protein-coupled receptor kinase 5"
/protein_id="CAI16487.1"
/db_xref="GI:55958736"
/db_xref="InterPro:IPR000239"
/db_xref="InterPro:IPR000342"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR008271"
/db_xref="InterPro:IPR011009"
/translation="MELENIVANTVLKAREGGGKRGKSKKWEILKPFPHISQCED
LRRTIDRYSGLDQKIGRLFRQFCETPGLECYIQFLDSVAEYVTPDEKLGEKG
KEIMTKYTLPKSPVFIAGQODLVISQTEELKLPCKELFSACAQSVEHYLRGPPEHE
YLDSEFFDRFLQWLKRPQVTKNTPROYVLGKGFGVCACOVRAIGKMYACKRLE
KKRIKKRGSMALEKLEKUNSFVNLAAYETKDALCLVLTIMGDLKFEHLY
NMGNPFEEALFYAAEILCGLEDHRENTVYRDLPENILDDYGHIRISDLGLAV
KIPGDDIRVRGVTGYMAPELNNQYGLSPDYWGICLIYEMIEGQSPFRGRKEV
KREEDVRVSYVHKFSEEAKSICMKLLTKDAQRKLCQCEGAAYKKRHPFFRN
MFKRLBAGMLDPPVPDRAVYCKGLVDIEQFSTVGKVLNHDTHDDDFYKFSFGSVS
IPQWNETECRELAVFGPNVTKLPDLNRNHPPEPPKGLLQLRKFQHQNNKSS
PSSKTSFNHINSHNSVNSITGSS"
join(73150. .73510,73847. .73941)
/gene="GPRK5"
/locus tag="RP11-714M16.1-003"
join(73150. .73510,73847. .73941)
/gene="GPRK5"
/locus tag="RP11-714M16.1-003"
/product="G protein-coupled receptor kinase 5"
/note="match: ESTs: AW170317.1"
75388
/gene="GPRK5"
/locus tag="RP11-714M16.1-002"
/note="Clone left_end: RP11-128N5"
76298. .76303
/gene="GPRK5"
/locus tag="RP11-714M16.1-002"
76321
/gene="GPRK5"
/locus tag="RP11-714M16.1-002"
/locus tag="RP11-714M16.1-002"
complement(join(120627. .120961,136211. .136354,
146734. .146820,148008. .148179))
/gene="RGS10"
/locus tag="RP11-714M16.2-002"
complement(join(120627. .120961,136211. .136354,
146734. .146820,148008. .148179))
/gene="RGS10"
/locus tag="RP11-714M16.2-002"
/product="regulator of G-protein signalling 10"
/note="match: ESTs: BQ890645.1 H87415.1
match: cDNAs: AF368902.1"
join(complement(AC012468.10:3196. .3314),
complement(148008. .148126),complement(146734. .146820),
complement(136211. .136354),complement(120627. .120961))
/gene="RGS10"
/locus tag="RP11-714M16.2-001"
join(complement(AC012468.10:3196. .3314),
complement(148008. .148126),complement(146734. .146820),
complement(136211. .136354),complement(120627. .120961))

/gene="RGS10"
/locus tag="RP11-714M16.2-001"
/product="regulator of G-protein signalling 10"
/note="match: ESTs: BG763741.1 BMD40804.1 BM914660.1
BM914667.1 BQ688292.1 BQ890645.1
match: cDNAs: BC009361.1"
complement(120627)
/gene="RGS10"
/locus tag="RP11-714M16.2-001"
complement(join(120782. .120961,136085. .136157,
146734. .146796))
/gene="RGS10"
/locus tag="RP11-714M16.2-004"
complement(join(120782. .120961,136085. .136157,
146734. .146796))
/gene="RGS10"
/locus tag="RP11-714M16.2-004"
/product="regulator of G-protein signalling 10"
/note="match: ESTs: BM838412.1"
complement(join(120815. .120961,136211. .136354,
146734. .146820,148008. .148151))
/gene="RGS10"
/locus tag="RP11-714M16.2-002"
/standard_name="OTTHUMP00000020598"
/note="match: proteins: O43665"
/codon_start=1
/product="regulator of G-protein signalling 10"
/protein_id="CAI16488.1"
/db_xref="GI:55958737"
/db_xref="InterPro:IPR000342"
/translation="MQSELCPADIHSDGSSSSHQSLKSTAKWAASLENLLEDEGV
KRREFPKKFSSENVLFWLACEDFKMQDKTQMQEKAKEIYMTFLSKASSQNVVEG
QSRNLKILPEHPLMFQKLQDQIFNLMKDYSRFLKSDLFLKHKRTTEEDLPDA
QTAAKRASRIYNT"
join(complement(AC012468.10:3196. .3244),
complement(148008. .148126),complement(146734. .146820),
complement(136211. .136354),complement(120815. .120961))
/gene="RGS10"
/locus tag="RP11-714M16.2-001"
/standard_name="OTTHUMP00000020597"
/note="match: proteins: Q96GN0"
/codon_start=1
/product="regulator of G-protein signalling 10"
/protein_id="CAI16486.1"
/db_xref="GI:55958735"
/db_xref="GOA:Q96GN0"

Query Match 90.6%; Score 15.4; DB 5; Length 153108;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACCCTTGCCCGAG 17
|||||
Db 90763 GCCCACCCTTGCCCAAG 90779

RESULT 75
AC150149/c
LOCUS
DEFINITION Gallus gallus clone WAG-78G23, WORKING DRAFT SEQUENCE, 15 ordered
pieces.
AC150149
AC150149.1 GI:49533744
VERSION HTG: HTGS PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 160714)
```

AUTHORS DOE Joint Genome Institute.
TITLE Unpublished
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 160714)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 2381390
Center clone name: Chk_78G23

Summary Statistics
Consensus quality: 153228 bases at least Q40
Consensus quality: 156522 bases at least Q30
Consensus quality: 158225 bases at least Q20
Estimated insert size: 191000; agarose-fp estimation
Estimated insert size: 159314; sum-of-contigs estimation
Quality coverage: 10.76 in Q20 bases; agarose-fp estimation
Quality coverage: 12.9 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 6200: contig of 6200 bp in length
* 6201 6300: gap of unknown length
* 6301 7878: contig of 1578 bp in length
* 7879 7878: gap of unknown length
* 7879 8896: contig of 918 bp in length
* 8897 8996: gap of unknown length
* 8997 19723: contig of 10726 bp in length
* 19723 19823: gap of unknown length
* 19823 45072: contig of 25250 bp in length
* 45073 45172: gap of unknown length
* 45173 46466: contig of 1293 bp in length
* 46466 46565: gap of unknown length
* 46566 48894: contig of 2329 bp in length
* 48895 48994: gap of unknown length
* 48995 54137: contig of 5143 bp in length
* 54138 54237: gap of unknown length
* 54238 69612: contig of 15375 bp in length
* 69613 69712: gap of unknown length
* 69713 82717: contig of 13004 bp in length
* 82717 82816: gap of unknown length
* 82817 118815: contig of 35999 bp in length
* 118816 118915: gap of unknown length
* 118916 132678: contig of 13763 bp in length
* 132679 132779: gap of unknown length
* 132780 139799: contig of 7020 bp in length
* 139799 139899: gap of unknown length
* 139899 155169: contig of 15270 bp in length
* 155169 155269: gap of unknown length
* 155269 160714: contig of 5446 bp in length.
Location/Qualifiers
1. 160714
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/clone_lib="WAG-78G23"
6201. .6300
/estimated_length=unknown

FEATURES
source
gap

gap 7879. .7978
/estimated_length=unknown
gap 8897. .8996
/estimated_length=unknown
gap 19723. .19822
/estimated_length=unknown
gap 45073. .45172
/estimated_length=unknown
gap 46466. .46565
/estimated_length=unknown
gap 48895. .48994
/estimated_length=unknown
gap 54138. .54237
/estimated_length=unknown
gap 59613. .69712
/estimated_length=unknown
gap 82717. .82816
/estimated_length=unknown
gap 118816. .118915
/estimated_length=unknown
gap 132679. .132778
/estimated_length=unknown
gap 139799. .139898
/estimated_length=unknown
gap 155169. .155268
/estimated_length=unknown

ORIGIN
Query Match 90.6%; Score 15.4; DB 12; Length 160714;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAG 17
|||||
DB 26346 GCCACGCTTGGCCGAG 26330
|||||

RESULT 76
AC011448 165122 bp DNA linear PRI 14-JUL-2002
LOCUS AC011448 Homo sapiens chromosome 19 clone CTC-260F20, complete sequence.
DEFINITION AC011448
ACCESSION AC011448
VERSION AC011448.5 GI:21747436
KEYWORDS HTG
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 165122)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 165122)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 165122)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 165122)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 165122)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

On Jul 14, 2002 this sequence version replaced gi:15281204.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www-shgc.stanford.edu
 Quality: Phrap Quality >=40 99.9% of Sequence;
 Estimated Total Number of Errors is 0.1.

FEATURES
 source
 Location/Qualifiers
 1. 165122
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTC-260P20"

ORIGIN
 Query Match 90.6%; Score 15.4; DB 5; Length 165122;
 Best Local Similarity 94.1%; Pred. No. 2.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17
 |||||
 Db 107847 GCCACGCTGGCCCAAG 107863

RESULT 77
 AC146997
 LOCUS AC146997 168122 bp DNA linear HTG 08-JUL-2004
 DEFINITION Arabcia punctulata clone BAP-144G2, WORKING DRAFT SEQUENCE, 13
 unordered pieces.
 AC146997
 AC146997.2 GI:50057837
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS Arabcia punctulata (punctuate urchin)
 SOURCE Arabcia punctulata
 ORGANISM Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidea; Euechinoidea; Echinacea; Arbacoidea; Arbacoidea;
 Arbacoidea.

REFERENCE
 AUTHORS
 1 (bases 1 to 168122)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Bmagu,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Briefe,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Garrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseg,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newcomb,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokweso,S., Ogum,H., Okwuonu,G.,
 Otaguaye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Fickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sison,I.,
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tanq,H.,

* 159006 159105: gap of unknown length
 * 159106 168122: contig of 9017 bp in length.
 FEATURES
 source

Location/Qualifiers
 1. .168122
 /organism="Arbacia punctulata"
 /mol_type="genomic DNA"
 /db_xref="taxon:7641"
 /clone="BAP-144G2"
 13949 .14048
 /estimated_length=unknown
 32580 .32679
 /estimated_length=unknown
 59557 .59656
 /estimated_length=unknown
 61694 .61793
 /estimated_length=unknown
 63798 .63897
 /estimated_length=unknown
 66963 .67062
 /estimated_length=unknown
 137627 .137726
 /estimated_length=unknown
 141732 .141831
 /estimated_length=unknown
 143877 .143976
 /estimated_length=unknown
 149391 .149490
 /estimated_length=unknown
 151665 .151764
 /estimated_length=unknown
 159006 .159105
 /estimated_length=unknown

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 168122;
 Best Local Similarity 94.1%; Pred. No. 2.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCCACGCTTGGCCGAG 17
 | ||||| ||||| |||||
 Db 80398 GACCACGCTTGGCCGAG 80414

RESULT 78
 AC130666/c
 LOCUS AC130666 172995 bp DNA linear ROD 29-SEP-2004
 DEFINITION Mus musculus chromosome 5, clone RP24-7907, complete sequence.
 ACCESSION AC130666
 VERSION AC130666.9 GI:52839751
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.

1 (bases 1 to 172995)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus chromosome 5, clone RP24-7907
 Unpublished
 2 (bases 1 to 172995)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Kamat,A., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,

Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNALREFERENCE
AUTHORS

Submitted (13-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 172995)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hagos,B., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Direct Submission

Submitted (06-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

4 (bases 1 to 172995)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hagos,B., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

COMMENT

Submitted (29-SEP-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 29, 2004 this sequence version replaced gi:50511630.

All repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WBIR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@broad.mit.edu

----- Project Information

Center project name: L27060

Center clone name: 79_O_7

Location/Qualifiers

1. .172995

FEATURES

source

Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

Location/Qualifiers

1..179959
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /map="q33.1-34"
 /clone="RP11-474D23"
 /clone_lib="RPCI-11.2"

misc_feature

1
 /note="Clone left end: RP11-474D23"

gene

join(4130..4224,4319..4410,11645..11755,23325..23394,26218..26393,27545..27709)

/gene="Cl3orf16"

mRNA

join(4130..4224,4319..4410,11645..11755,23325..23394,26218..26393,27545..27709)

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-002"

/product="chromosome 13 open reading frame 16"

/note="match: ESTs: AI004422.1 AW070313.1 AW665078.1

BF509929.1 BI828648.1 BI830521.1

match: cDNA8: BC029889.1"

join(AL139086.12:84885..84958,4319..4410,11645..11755,22756..22900,23325..23394,26218..26393,27545..27711)

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-001"

join(AL139086.12:84885..84958,4319..4410,11645..11755,22756..22900,23325..23394,26218..26393,27545..27711)

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-001"

/product="chromosome 13 open reading frame 16"

/note="match: ESTs: AA369808.1 AA470135.1 AW957098.1"

join(AL139086.12:84912..84958,4319..4410,11645..11755,22756..22760)

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-001"

/standard_name="OTTHUMP0000018719"

/codon_start=1

/product="chromosome 13 open reading frame 16"

/protein_id="CAH74151.1"

/db_xref="GI:55662217"

/db_xref="InterPro:IPR000519"

/db_xref="UniProtKB/TrEMBL:Q5VZG8"

/translation="MPATICQGRQVTVQACARPLHLSAAAMEYVLEVGNSPRHLLKQFTVCDVPLYDICTNVNSDRRCQELGCGCEGVCKKAVPI"

join(4353..4410,11645..11755,23325..23394,26218..26393,27545..27585)

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-002"

/standard_name="OTTHUMP0000018720"

/codon_start=1

/product="chromosome 13 open reading frame 16"

/protein_id="CAH74150.1"

/db_xref="GI:55662216"

/db_xref="InterPro:IPR000519"

/db_xref="UniProtKB/TrEMBL:Q8N6K0"

/translation="MEYVLEVKNSPRHLLKQFTVCDVPLYDICTNVNSDRRCQELGCGCEGVCKKAVPIIHVPSALIVIIAGAFVITIIYRVIOESRKEKALPDVVALPKQSSFYEGVCYKXAVPI"

EKAELASSKGLGKLPASPGPPSAGPSNKSDEKDDVTGTTTEAETED"

13804..14058

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-001"

polyA_signal

27689..27694

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-001"

polyA_site

27709

/gene="Cl3orf16"

/locus_tag="RP11-474D23.1-001"

misc_feature

37339

/note="Clone right end: RP11-494P5"

40640..41524

/note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."

misc_feature

41892..41696

/note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."

misc_feature

59653..59994

/note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."

misc_feature

105916

/note="Clone left end: RP11-190K22"

misc_feature

118600

/note="Tandem repeat. Forced join. Gap size estimated to be approximately 500bp by restriction digest data."

misc_feature

176889..176929

/note="Sequence from overlapping clone BA190K22 (AL162714). Assembly confirmed by restriction digest data."

misc_feature

179959

/note="Clone right end: RP11-474D23"

ORIGIN

Query Match 90.6%; Score 15.4; DB 5; Length 179959;

Best Local Similarity 94.1%; Pred. No. 2.3e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

|||||

Db 145314 GCCCAGCGTTGGCCGAG 145298

|||||

RESULT 80

AC017033/c

LOCUS AC017033 183887 bp DNA linear PRI 21-APR-2005

DEFINITION Homo sapiens BAC clone RP11-297J22 from 2, complete sequence.

ACCESSION AC017033

VERSION AC017033.5 GI:12740494

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo

REFERENCE 1 (bases 1 to 183887)

Maupin,R., Kruchowski,S. and Haakenson,W.

The sequence of Homo sapiens BAC clone RP11-297J22

Unpublished (2001)

REFERENCE 2 (bases 1 to 183887)

Waterston,R.H.

Direct Submission

Submitted (09-DEC-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

REFERENCE 3 (bases 1 to 183887)

Waterston,R.H.

Direct Submission

Submitted (10-FEB-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

REFERENCE 4 (bases 1 to 183887)

```

AUTHORS
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS 5 (bases 1 to 183887)
TITLE Waterston, R.
JOURNAL Direct Submission
Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 183887)
AUTHORS Waterston, R.
TITLE Direct Submission
Submitted (10-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 183887)
AUTHORS Wilson, R.K.
TITLE Direct Submission
Submitted (21-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Feb 10, 2001 this sequence version replaced gi:10999016.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics
----- Center project name: H_NH0297J22
-----

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-57J18; the clone sequenced
to the right is RP11-549G13. Actual start of this clone is at base
position 1 of RP11-297J22; actual end is at base position 183887 of
RP11-297J22.

RP11-297J22 contains a tandem repeat between bases 149996 to 150321
where there are low quality bases. Digest and pcr information
confirm the assembly.

There are polymorphisms between RP11-297J22 and RP11-549G13.
Location/Qualifiers
1..183887
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
/chromosome="2"
/clone="RP11-297J22"
/clone_lib="RPCI-11"
3664..6040
/misc_feature
/note="CpG island (GC=67.4, o/e=0.81, #CpGs=212)"
9984..10384
/misc_feature
/note="CpG island (GC=77.1, o/e=1.02, #CpGs=61)"
60665..65850
/gene="GLI2"
join(60665..60820,65673..65850)
/gene="GLI2"
/product="unknown"
/note="Homo sapiens GLI-Kruppel family member GLI2 (GLI2),
transcript variant 1, mRNA.; H_NH0297J22.1
This gene was based on gi(13518230)
Continues as H_NH0549G13.1"
135634..136015
/misc_feature
/note="CpG island (GC=68.6, o/e=0.67, #CpGs=34)"

ORIGIN
Query Match 90.6%; Score 15.4; DB 5; Length 183887;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCGGAG 17
Db 4738 GCCGCGCTGGCGGAG 4722

RESULT 81
AC019078
LOCUS AC019078 187027 bp DNA linear HTG 30-NOV-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-330P9, WORKING DRAFT SEQUENCE,
22 unordered pieces.
ACCESSION AC019078
VERSION AC019078.6 GI:11465145
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE HTG; Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 187027)
Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187027)
Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Nov 30, 2000 this sequence version replaced gi:9887808.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0330P09
----- Summary Statistics -----
Sequencing vector: M13; 9%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 9% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171161 bases at least Q40
Consensus quality: 176449 bases at least Q30
Consensus quality: 179564 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 184927; sum-of-contigs
Quality coverage: 4.11 in Q20 bases; agarose-fp
Quality coverage: 4.12 in Q20 bases; sum-of-contigs

```

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 22 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 1119: contig of 1119 bp in length
* 1120 1219: gap of unknown length
* 1220 2642: contig of 1423 bp in length
* 2643 2742: gap of unknown length
* 2743 7458: contig of 4716 bp in length
* 7459 7559: gap of unknown length
* 11010: contig of 3452 bp in length
* 11011: gap of unknown length
* 11111 15060: contig of 3950 bp in length
* 15061 15160: gap of unknown length
* 15161 19635: contig of 4475 bp in length
* 19636 19735: gap of unknown length
* 19736 24016: contig of 4281 bp in length
* 24017 24117: gap of unknown length
* 24117 29667: contig of 5551 bp in length
* 29668 35444: gap of unknown length
* 35445 40954: contig of 5410 bp in length
* 40955 48233: contig of 7179 bp in length
* 48234 48334: gap of unknown length
* 48334 57406: contig of 9073 bp in length
* 57407 57506: gap of unknown length
* 57507 66359: contig of 8853 bp in length
* 66360 76814: contig of 10355 bp in length
* 76815 76914: gap of unknown length
* 76915 87795: contig of 10781 bp in length
* 87796 98686: contig of 10891 bp in length
* 98687 112322: contig of 13536 bp in length
* 112323 124223: gap of unknown length
* 124223 126002: contig of 13580 bp in length
* 126003 126102: gap of unknown length
* 126103 139276: contig of 13174 bp in length
* 139277 139376: gap of unknown length
* 139377 159274: contig of 19898 bp in length
* 159275 159374: gap of unknown length
* 159375 185651: contig of 26277 bp in length
* 185652 185751: gap of unknown length
* 185752 187027: contig of 1276 bp in length.
```

FEATURES

source

```

1..187027
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="2"
  /clone="RP11-330P9"
```

misc_feature

```
1..1119
  /note="assembly_name:Contig10"
```

gap

```
1120..1219
  /estimated_length=unknown
```

misc_feature

```
1220..2642
  /note="assembly_name:Contig13"
```

gap

```
2643..2742
  /estimated_length=unknown
```

misc_feature

```
2743..7458
  /note="assembly_name:Contig14"
```

gap

```
7459..7558
  /estimated_length=unknown
```

misc_feature

```
7559..11010
  /note="assembly_name:Contig15"
```

```

gap
  /estimated_length=unknown
misc_feature
1111..15060
  /note="assembly_name:Contig16"
15061..15160
  /estimated_length=unknown
15161..19635
  /note="assembly_name:Contig17"
19636..19735
  /estimated_length=unknown
19736..24016
  /note="assembly_name:Contig18"
24017..24116
  /estimated_length=unknown
24117..29667
  /note="assembly_name:Contig19"
29668..29767
  /estimated_length=unknown
29768..35444
  /note="assembly_name:Contig20"
35445..35544
  /estimated_length=unknown
35545..40954
  /note="assembly_name:Contig21"
40955..41054
  /estimated_length=unknown
41055..48233
  /note="assembly_name:Contig22"
48234..48333
  /estimated_length=unknown
48334..57406
  /note="assembly_name:Contig23"
57407..57506
  /estimated_length=unknown
57507..66359
  /note="assembly_name:Contig24"
66360..66459
  /estimated_length=unknown
66460..76814
  /note="assembly_name:Contig25"
76815..76914
  /estimated_length=unknown
76915..87795
  /note="assembly_name:Contig26"
87796..87795
  /estimated_length=unknown
87796..98686
  /note="assembly_name:Contig27"
98687..98786
  /estimated_length=unknown
98787..112322
  /note="assembly_name:Contig28"
112323..112422
  /estimated_length=unknown
112423..126002
  /note="assembly_name:Contig29"
126003..126102
  /estimated_length=unknown
126103..139276
  /note="assembly_name:Contig30"
139277..139376
  /estimated_length=unknown
139377..159274
  /note="assembly_name:Contig31"
159275..159374
  /estimated_length=unknown
159375..185651
  /note="assembly_name:Contig32"
185652..185751
  /estimated_length=unknown
185752..187027
  /note="assembly_name:Contig9"
```

ORIGIN

```
Query Match      90.6%; Score 15.4; DB 12; Length 187027;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
Db 154552 GCCCGCGCTGGCCGAG 154568

RESULT 82
AC150074/c
LOCUS      AC150074
DEFINITION Gallus gallus clone CH261-96B15, WORKING DRAFT SEQUENCE, 9 ordered
            pieces.
ACCESSION   AC150074
VERSION     AC150074.1 GI:49533669
KEYWORDS    HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ACTIVEFIN.
SOURCE      Gallus gallus (chicken)
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 193015)
            DOE Joint Genome Institute.
            Unpublished
            Unpublished
REFERENCE   2 (bases 1 to 193015)
            DOE Joint Genome Institute.
            Direct Submission.
            Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
            94598-1698, USA
COMMENT     -----
            Center: Joint Genome Center
            Center Code: JGI
            Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 2967678
Center clone name: JF2-96B15
-----
Summary Statistics
Consensus quality: 189480 bases at least Q40
Consensus quality: 190386 bases at least Q30
Consensus quality: 191284 bases at least Q20
Estimated insert size: 204000; agarose-fp estimation
Estimated insert size: 192215; sum-of-contigs estimation
Quality coverage: 9.75 in Q20 bases; agarose-fp estimation
Quality coverage: 10.35 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 74911 75010: contig of 74910 bp in length
* 75011 79172: contig of 4162 bp in length
* 79173 79272: gap of unknown length
* 79273 89913: contig of 10641 bp in length
* 89914 90013: gap of unknown length
* 90014 104546: contig of 14533 bp in length
* 104547 104646: gap of unknown length
* 104647 112802: contig of 8156 bp in length
* 112803 112902: gap of unknown length
* 112903 130534: contig of 17632 bp in length
* 130535 130634: gap of unknown length
* 130635 133851: contig of 3217 bp in length
* 133852 133951: gap of unknown length
```

```
* 133952 136970: contig of 3019 bp in length
* 136971 137070: gap of unknown length
* 137071 193015: contig of 55945 bp in length.
FEATURES             source
     1..193015
     /organism="Gallus gallus"
     /mol_type="genomic DNA"
     /db_xref="taxon:9031"
     /clone_lib="CH261-96B15"
     /clone_lib="CHORI-261 Chicken BAC Library"
     74911..75010
     /estimated_length=unknown
     79173..79272
     /estimated_length=unknown
     89914..90013
     /estimated_length=unknown
     104547..104646
     /estimated_length=unknown
     112803..112902
     /estimated_length=unknown
     130535..130634
     /estimated_length=unknown
     133852..133951
     /estimated_length=unknown
     136971..137070
     /estimated_length=unknown

ORIGIN
Query Match      90.6%; Score 15.4; DB 12; Length 193015;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
Db 111419 GCCCAGCGCTGGCCGAG 111403

RESULT 83
AX417035
LOCUS      AX417035
DEFINITION Sequence 4026 from Patent WO0228891.
ACCESSION   AX417035
VERSION     AX417035.1 GI:21449645
KEYWORDS    Listeria innocua
SOURCE      Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
            Kunst, F. and Glaser, P.
            Listeria innocua genome and applications
            Patent: WO 0228891-A 4026 11-APR-2002;
            INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE
            SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1..195269
     /organism="Listeria innocua"
     /mol_type="unassigned DNA"
     /db_xref="taxon:1642"
     /note="second part of seq 8
            first part = seq 8 (000001 - 349980)
            second part = seq 4026 (300001 - 495269)"

ORIGIN
Query Match      90.6%; Score 15.4; DB 2; Length 195269;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
Db 100813 GCCCAGCGCTGGCTGAG 100829

RESULT 84
```



```

AC002365/c
LOCUS       AC002365               196542 bp    DNA    linear    PRI 22-AUG-1998
DEFINITION  Homo sapiens chromosome X clone U177G4, U152H5, U168D5, 174A6,
U172D6, and U186B3 from Xp22, complete sequence.
ACCESSION   AC002365 U82630 U83511
VERSION     AC002365.1 GI:258015
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 196542)
Muzny,D., Aronson,A.D., Adams,C., Brundage,E., Bunac,C.,
Carvelli,K., Chacko,J., Chen,J., Di W., Ding,Y., Dugan,S.,
Durbin,J., Forcum,J., Ganesh,R., Garcia,C., Goodman,M.,
Gorrell,J.H., Haywood,M., Hernandez,J., Jackson,L., Jin,S.,
Kampal,R., Karpthy,S., Kovar,C., Leal,B., Li,Y., Lichtarge,O.,
Liu,W., Logan,O., Lu,J., Ly,T., Martinez,C., Oswal,G., Perez,L.,
Rashid,N.D., Rowland,K., Savage,L., Scherer,S.E., Shen,H.,
Simon,M., Stovall,K., Timms,K.M., Todd,J., Vo,Q., Williamson,A.,
Worley,K.C., Yu,W., Chinault,C., Nelson,D. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 196542)
Chiu,M.W.
Direct Submission
Submitted (23-JUL-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 196542)
Muzny,D., Aronson,A.D., Adams,C., Brundage,E., Bunac,C., Lu,J.,
Shen,Y., Rowland,K., Liu,W., Perez,L., Ding,Y., Gonzalez,O.,
Haywood,M., Jain,A., Leal,B., Logan,O., Nguyen,V., Savage,L.,
Shen,H., Worley,K., Chen,E., Forcum,J., Aronson,A.D., Chiu,M.W.,
Gorrell,J.H., Brundage,E., Di W., Chinault,C., Nelson,D. and
Gibbs,R.A.
Direct Submission
Submitted (20-AUG-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 196542)
Chiu,M.W.
Direct Submission
Submitted (04-SEP-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA
5 (bases 1 to 196542)
Worley,K.C.
Direct Submission
Submitted (22-AUG-1998) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On or before Sep 4, 1997 this sequence version replaced gi:1766071,
gi:177380, gi:2289024.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
Zhang.

Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.
Location/Qualifiers
1. 196542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"

FEATURES             source

```

```

/map="Xp22"
/clone="U177G4"
/LONGC01"
complement(760..1057)
/rpt_family="AluJb"
1893..2194
/rpt_family="AluSg"
2305..2584
/rpt_family="AluSx"
3418..3630
/rpt_family="MER20"
complement(3806..4096)
/rpt_family="AluSx"
4138..4537
/rpt_family="L1MB2"
complement(4549..4841)
/rpt_family="AluJb"
complement(5421..5528)
/rpt_family="MER58B"
complement(7001..7132)
/rpt_family="FAM"
7773..7903
/rpt_family="AluJ"
7910..8215
/rpt_family="AluYa5"
8224..8401
/rpt_family="AluJ"
complement(8932..9232)
/rpt_family="AluY"
complement(9458..9481)
/rpt_family="POLY_A"
11579..11618
/rpt_family="AT rich"
complement(11619..11916)
/rpt_family="AluSg"
complement(11917..12252)
/rpt_family="AluJb"
complement(12328..12633)
/rpt_family="AluSx"
complement(12665..13117)
/rpt_family="L1MB3"
complement(13236..13267)
/rpt_family="L1MA1"
13323..13604
/rpt_family="AluJb"
13983..14378
/notes="Similar to sequence with GenBank Accession number
Z36840"
15384..15565
/rpt_family="MER58B"
15856..15956
/rpt_family="AluSg"
16091..16660
/rpt_family="L1MB7"
<17023..>92814
/gene="HSAPXL"
join(<17023..17176,34350..34481,37731..40071,41563..41663,
75348..76243,80507..81058,82568..82739,88014..88286,
90044..>92814)
/gene="HSAPXL"
join(<17023..17176,34350..34481,37731..40071,41563..41663,
75348..76243,80507..81058,82568..82739,88014..88286,
90044..>92814)
/gene="HSAPXL"
/product="APXL"
join(<17023..17176,34350..34481,37731..40071,41563..41663,
75348..76243,80507..81058,82568..82739,88014..88286,
90044..90310)
/gene="HSAPXL"
/notes="human homolog of Xenopus laevis APX gene, similar
to sequence encoded by GenBank Accession Number X83543"
/codon_start=3

```

```
/product="APXL"
/protein_id="AAC32592.1"
/db_xref="GI:1773381"
/translation="IEGSKAAAVDKLLAGDEIVGINDIGLSGFPROBAICLVKGSHTK
LKLVKRSSELGWRPSPHSHATKESDHPPELAASPFTSTSCPSWGRHSHASSSHDL
SSWEQTLQRTLHSSLSGSDSLDHPSSRLSVAKSNSSIDHLGSHKRSASVGSFST
GDSKGRPRPFAEPAKAPGRSNFVWVYDPPKAPSPPPPPPLRSDSFAATKS
HEKAQGVFSEAAAQHTALAQAPGRDRPELTDPRMSAHPGSLGKSGGPGCPQ
EAAHAGSWPSPKQSRGSLSSSDVFPQSPHSGRHPPLYSNDSPLCADSLGQBP
GAASFONDSPOVGLSSCDQKLGSGWQRPQVQDLQAQLWAGCWPSDTALGALE
SLPPTVQGPRLHLPQEGPPDARETRGCYPLDKGAEGSAGAQEPFRASRAEKASQ
RLAASITWAGESSRICQETPLLSLUTQEGKRPESSPDSATRPFPFPAHVCKPTR
RSRFAITLNEIGSMRAIKQKSTVALTAAGEAEDGTGRWAGLGGVQEGPLACT
YKDLHKAQARVLRRTTAQKLSYSEPEKMEVGLTRGYSQHQPRTSEDTVGTFA
STGGPHPRGRRTTAQKLSYSEPEKMEVGLTRGYSQHQPRTSEDTVGTFA
DRMKFEETSKVPQPAQKQALHGIIPDKPERPRTAGRTCEGTEPWSRTTSLGDSLN
AHSAAEAGTSIDLPRRLGTAEVQASWKEQRKPLEARSSGRCHSADDILOVSLDPQR
FQVHGRSRSPSTDIHYQASVELRQAGDPGEFBEELFSAVRAEAGQSTPQADAO
CREGSPGQHPSPKAPNPPTFSELSCRGAPELPREGRAGTLPRDRIYSEESTP
ADLGPRAQSPGSLHARGQDSWPVSSALLSKRPAPQPPPPKRPRIYRATDGPADA
PVGVLGRPFTPTPASLDVVARLSLHSPSVFSAQFQDTPKATVCERGSHVSGDA
SRPLPEALLPKQOHLRLQATMETSPSPSPQAPQKLTDPKPLLIODESRTIERVM
DNNTYKXVIKIVHSSQPEKESRQSLACPAEPALPHGLEQDKIITLSTSEQFYR
FCLYTRQGAEPAPQAPAEPLQTVPPKRCCTSPGSLSYKAKKEKTVEDLKS
ELAREIVGDKSLADILDPVKIKTMDLMEGIPKDEHLLLEEAQRRKLLPKIPSPR
STERKEEPSVPAASVATNSTYTSAPKAELLIKMDLQEQHEEDSGSDLDHDL
SVKKQELIESTSRKQVLRARESLLEDVQANTVIGAEVAIKVGCCKPSEFQKRFM
IGDLKDVNLLLSLGLARVENALNLDGASPGDRQSLEKQVLIQOHEDAKELK
ENLDRRRIIVPDILANLYLSLESADYEHFVYKMSALLIEQRELEDKIHLGEEQKCLL
DSLQPERGK"
repeat_region complement(17976. .18074)
/rpt_family="HY3"
repeat_region 18162. .18473
/rpt_family="AluJb"
repeat_region 19636. .19717
/rpt_family="MER21A"
repeat_region complement(19798. .19937)
/rpt_family="L1MB3"
repeat_region complement(20181. .20327)
/rpt_family="FRAM"
repeat_region complement(22367. .22401)
/rpt_family="Alu-rich"
repeat_region 22406. .22461
/rpt_family="(GGAA)n"
repeat_region complement(22462. .22627)
/rpt_family="AluJo"
repeat_region complement(22638. .22933)
/rpt_family="AluX"
repeat_region complement(26881. .26957)
/rpt_family="MERSA"
repeat_region complement(27726. .27762)
/rpt_family="POLY A"
repeat_region complement(27876. .27977)
/rpt_family="AluJo/FRAM"
```

```
Query Match 90.6%; Score 15.4; DB 5; Length 196542;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
|||||
Db 17324 GCCATGCTTGCCGAG 17308

RESULT 85
AC005054/c AC005054 197456 bp DNA linear HTG 17-AUG-2000
LOCUS Homo sapiens chromosome UNK clone CTB-45HS, *** SEQUENCING IN
PROGRESS ***, 18 unordered pieces.
ACCESSION AC005054
VERSION AC005054.2 GI:9838352
KEYWORDS HTG; HTGS_PHASE1.
```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 197456)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 197456)
Waterston,R.H.
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 17, 2000 this sequence version replaced gi:3212937.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H RG045H05
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 71% of reads
Chemistry: Dye-terminator Big Dye; 29% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: bases at least Q40
Consensus quality: bases at least Q30
Consensus quality: bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 195258; sum-of-contigs
Quality coverage: in Q20 bases; agarose-fp
Quality coverage: in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 5084: contig of 5084 bp in length
* 5085 5184: gap of unknown length
* 5185 14399: contig of 9215 bp in length
* 14400 14499: gap of unknown length
* 14500 22837 22936: contig of 8337 bp in length
* 22937 33373: contig of unknown length
* 33374 33473: contig of 10437 bp in length
* 33474 47419: contig of 13946 bp in length
* 47420 47519: gap of unknown length
* 47520 66575: contig of 19056 bp in length
* 66576 81990: contig of 15315 bp in length
* 81991 82091: gap of unknown length
* 82091 105847: contig of 23757 bp in length
* 105848 129271: contig of 23324 bp in length
* 129272 129371: gap of unknown length
* 129372 164242: contig of 34871 bp in length
* 164243 164342: gap of unknown length
* 164343 164840: contig of 498 bp in length
* 164841 164940: gap of unknown length
* 164941 167875: contig of 2935 bp in length
* 167876 167975: gap of unknown length
* 167976 172505: contig of 4530 bp in length
* 172506 172605: gap of unknown length
* 172606 176804: contig of 4199 bp in length
* 176805 176904: gap of unknown length
```

* 176905 180960: contig of 4056 bp in length
 * 180961 181060: gap of unknown length
 * 181061 185808: contig of 4748 bp in length
 * 185809 185908: gap of unknown length
 * 185909 190832: contig of 4924 bp in length
 * 190833 190932: gap of unknown length
 * 190933 197456: contig of 6524 bp in length.

FEATURES

source Location/Qualifiers

1. 197456
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="UNK"
 /clone="CTB-45H5"
 1. 5084
 /note="assembly_name:Contig100"
 5085..5184
 /estimated_length=unknown
 5185..14339
 /note="assembly_name:Contig101"
 14400..14499
 /estimated_length=unknown
 14500..22836
 /note="assembly_name:Contig103"
 22837..22936
 /estimated_length=unknown
 22937..33373
 /note="assembly_name:Contig104"
 33374..33473
 /estimated_length=unknown
 33474..47419
 /note="assembly_name:Contig105"
 47420..47519
 /estimated_length=unknown
 47520..66575
 /note="assembly_name:Contig106"
 66576..66675
 /estimated_length=unknown
 66676..81390
 /note="assembly_name:Contig107"
 81391..82090
 /estimated_length=unknown
 82091..105847
 /note="assembly_name:Contig108"
 105848..105947
 /estimated_length=unknown
 105948..129271
 /note="assembly_name:Contig109"
 129272..129371
 /estimated_length=unknown
 129372..164242
 /note="assembly_name:Contig110"
 164243..164342
 /estimated_length=unknown
 164343..164840
 /note="assembly_name:Contig18"
 164841..164940
 /estimated_length=unknown
 164941..167875
 /note="assembly_name:Contig91"
 167876..167975
 /estimated_length=unknown
 167976..172505
 /note="assembly_name:Contig94"
 172506..172605
 /estimated_length=unknown
 172606..176804
 /note="assembly_name:Contig95"
 176805..176904
 /estimated_length=unknown
 176905..180960
 /note="assembly_name:Contig96"
 180961..181060

misc_feature
 181061..185808
 /note="assembly_name:Contig97"
 gap
 185809..185908
 /estimated_length=unknown
 misc_feature
 185909..190832
 /note="assembly_name:Contig98"
 gap
 190833..190932
 /estimated_length=unknown
 misc_feature
 190933..197456
 /note="assembly_name:Contig99"

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 197456;
 Best Local Similarity 94.1%; Pred. No. 2.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCCGAG 17
 |||||
 Db 131397 GCCACGCTTGGCCGAG 131381

RESULT 86

AC120873/c

LOCUS 198375 bp DNA linear ROD 15-JUN-2004
 DEFINITION Mus musculus chromosome 12, clone RP23-283P14, complete sequence.

AC120873

AC120873.7 GI:48717694

HTG.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

Birren,B., Nusbaum,C. and Lander,E.

1 (bases 1 to 198375)

Mus musculus chromosome 12, clone RP23-283P14

Unpublished

2 (bases 1 to 198375)

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gordon,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,P., Roy,A., Santos,R., Schauer,S., Schupack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

```

/rpt_family="TTTTGG)n"
complement(3311..3997)
/rpt_family="L1_MM"
4329..4419
/rpt_family="(TCCC)n"
complement(4793..4956)
/rpt_family="MER33"
5201..5329
/rpt_family="B1_MM"
5335..5368
/rpt_family="(CCAA)n"
6843..6888
/rpt_family="(TTTG)n"
complement(6890..6970)
/rpt_family="B1_MM"
6974..7278
/rpt_family="L1M4"
7654..7776
/rpt_family="L1M4"
7884..8040
/rpt_family="RSINE1"
8524..8623
/rpt_family="Tigger7"
8614..8811
/rpt_family="Tigger7"
8881..9168
/rpt_family="B4"
complement(9304..9545)
/rpt_family="B4A"
9848..9874
/rpt_family="AT_rich"
10732..10800
/rpt_family="(TG)n"
complement(13449..14096)
/rpt_family="Lx6"
14267..14443
/rpt_family="B2_Mml1"
15144..15202
/rpt_family="ID_B1"
15428..16421
/rpt_family="L1_MM"
17126..17738
/rpt_family="AT_rich"
complement(17998..18401)
/rpt_family="MTB"
18759..18795
/rpt_family="(TCTG)n"
complement(18951..19258)
/rpt_family="MTD"
19369..19779
/rpt_family="MTC"
20027..20317
/rpt_family="B4"
complement(21317..21467)
/rpt_family="L2"
23420..23588
/rpt_family="(TG)n"
24688..24724
/rpt_family="AT_rich"
25083..25161
/rpt_family="MSTC"
complement(25873..27111)
/rpt_family="L1_MM"
27105..27498
/rpt_family="L1"
28194..28226
/rpt_family="(TTTA)n"
28232..28257
/rpt_family="(T)n"
complement(28258..28404)
/rpt_family="B1_MM"
complement(29655..30638)
/rpt_family="Lx2"

```

```

repeat_region      complement(30641. .30804)
repeat_region      /rpt_family="Lx2"
repeat_region      complement(30847. .30950)
repeat_region      /rpt_family="Lx2"
repeat_region      complement(31085. .32311)

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 6; Length 198375;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
|||
Db 23971 GCACACGCTTGCCGAG 23955

RESULT 87
AC009061 AC009061 200618 bp DNA linear PRI 26-JUN-2001
DEFINITION Homo sapiens chromosome 16 clone RP11-297D21, complete sequence.
ACCESSION AC009061
VERSION AC009061.10 GI:14550301
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Hominidae; Homo
REFERENCE 1 (bases 1 to 200618)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 200618)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 200618)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jun 26, 2001 this sequence version replaced gi.9954595.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.4.
STS Content:
SHGC-13828 G14855
SHGC-33002 G27705
FEATURES
source Location/Qualifiers
1. .200618
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-297D21"

ORIGIN
Query Match 90.6%; Score 15.4; DB 5; Length 200618;
Best Local Similarity 94.1%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
|||
Db 94450 GCCACGCTTGCCGAG 94466

RESULT 88
AC119828 AC119828 205915 bp DNA linear ROD 05-NOV-2003
LOCUS AC119828
DEFINITION Mus musculus chromosome 10, clone RP23-186A24, complete sequence.

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC119828
AC119828.9 GI:38175966
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Scuriognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 205915)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 10, clone RP23-186A24
Unpublished
2 (bases 1 to 205915)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campioiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 205915)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McEwan,P., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (11-OCT-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 205915)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,

Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ranasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (05-NOV-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 5, 2003 this sequence version replaced gi:37620325.
 All repeats were identified using RepeatMasker:
 Snit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19587

Center clone name: 186_A_24

FEATURES

source

Location/Qualifiers

1..205915

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="10"

/map="10"

/clone_lib="RP23-186A24"

/clone="RP23-186A24"

1..2187

misc_feature

/note="wgs end extension"

clone_end:SP6

unsure

/note="complement(34..38)"

/note="<30 qual SNGL region"

complement(92..96)

/note="<30 qual SNGL region"

complement(906..1320)

/rpt_family="MTB"

1326..1424

/rpt_family="B4A"

1804..1850

/rpt_family="(CCA)n"

1917..1946

/rpt_family="(CAGA)n"

2188..2193

/note="clone boundary"

clone_end:SP6

site:EcoRI

2632..2723

/rpt_family="L2"

complement(3241..3331)

/rpt_family="MIR"

4268..4292

/rpt_family="(CA)n"

4510..4688

/rpt_family="L1ME"

5052..5421

/rpt_family="Lx8"

5435..5457

/rpt_family="(TTTG)n"

5467..5810

/rpt_family="Lx8"

5884..5959

/rpt_family="RMER15"

6254..6276

/rpt_family="AT_rich"

6278..6427

repeat_region

repeat_region
 /rpt_family="B4A"
 6599..6718
 /rpt_family="MER5A"
 7692..7968
 /rpt_family="B4"
 10091..10172
 /rpt_family="RLTR14"
 11372..11508
 /rpt_family="B1_MM"
 11509..11535
 /rpt_family="B4"
 12285..12405
 /rpt_family="L1MA4A"
 12591..12643
 /rpt_family="CT-rich"
 12789..12901
 /rpt_family="L1MA4A"
 12941..12966
 /rpt_family="(TTA)n"
 12981..13045
 /rpt_family="(TG)n"
 complement(13048..13169)
 /rpt_family="RSINE1"
 13170..13311
 /rpt_family="L1MA4A"
 13523..13549
 /rpt_family="AT_rich"
 complement(14629..14807)
 /rpt_family="URR1B"
 14969..15011
 /rpt_family="(TG)n"
 complement(15416..15506)
 /rpt_family="MIR"
 15658..15680
 /rpt_family="(CAGA)n"
 15940..16034
 /rpt_family="Lx2"
 16564..16926
 /rpt_family="ORR1B2"
 complement(17275..17583)
 /rpt_family="ORR1A3"
 complement(18440..18617)
 /rpt_family="B4A"
 complement(18632..18834)
 /rpt_family="URR1A"
 19286..19807
 /rpt_family="L1MB8"
 19808..20133
 /rpt_family="ORR1A2"
 20134..20527
 /rpt_family="L1MB8"
 20884..20913
 repeat_region

Query Match 90.6%; Score 15.4; DB 6; Length 205915;

Best Local Similarity 94.1%; Pred. No. 2.4e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17

|||||

Db 199705 GCCCAGCTTGCCCAAG 199721

RESULT 89

AC021091/c

LOCUS

DEFINITION

AC021091

SEQUENCE, 8 ordered pieces.

AC021091

VERSION

AC021091.5

KEYWORDS

HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AC021091 Homo sapiens chromosome 16 clone RP11-421N4, WORKING DRAFT

209262 bp DNA linear HTG 25-APR-2001

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 209262)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 209262)
DOE Joint Genome Institute.
Direct Submission
Submitted (14-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 25, 2001 this sequence version replaced gi:9954658.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Project Information
Center Project Name: 568345
Center clone name: RPCI-11_421N4

Summary Statistics
Consensus quality: 201548 bases at least Q40
Consensus quality: 207017 bases at least Q30
Consensus quality: 207950 bases at least Q20
Estimated insert size: 238990; agarose-fp estimation
Estimated insert size: 208912; sum-of-contigs estimation
Quality coverage: 6.28 in Q20 bases; agarose-fp estimation
Quality coverage: 7.19 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 5137: contig of 5137 bp in length
* 5138 5237: gap of unknown length
* 5238 47243: contig of 42006 bp in length
* 47244 47343: gap of unknown length
* 47344 68342: contig of 20999 bp in length
* 68343 68442: gap of unknown length
* 68443 151499: contig of 83057 bp in length
* 151500 151599: gap of unknown length
* 151600 161329: contig of 9730 bp in length
* 161330 163824: contig of 2395 bp in length
* 161430 163825: gap of unknown length
* 163825 163925: gap of unknown length
* 163925 202407: contig of 38483 bp in length
* 202408 202507: gap of unknown length
* 202508 209262: contig of 6755 bp in length.
Location/Qualifiers
1. 209262
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-421N4"
/clone_lib="RPCI human BAC library 11"
5138..5237
/estimated_length=unknown
47244..47343
/estimated_length=unknown
68343..68442
/estimated_length=unknown
151500..151599
/estimated_length=unknown
161330..161429
/estimated_length=unknown
163825..163924
/estimated_length=unknown

FEATURES
source

1. 209262
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-421N4"
/clone_lib="RPCI human BAC library 11"
5138..5237
/estimated_length=unknown
47244..47343
/estimated_length=unknown
68343..68442
/estimated_length=unknown
151500..151599
/estimated_length=unknown
161330..161429
/estimated_length=unknown
163825..163924
/estimated_length=unknown

gap
ORIGIN
Query Match 90.6%; Score 15.4; DB 12; Length 209262;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCCAGCGCTGGCCGAG 17
|||||||
DB 80342 GCCCAGCGCTGGCCGAG 80326
RESULT 90
AC091522 214798 bp DNA linear ROD 27-NOV-2002
LOCUS Mus musculus strain C57BL/6J chromosome 10 clone rp23-199g2,
DEFINITION complete sequence.
AC091522
AC091522 AC091522.23 GI:25703607
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE Mus musculus Chromosome 10 BAC Clone rp23-199g2
JOURNAL Unpublished
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL
Ying, N., Song, L. and Roe, B.A.
2 (bases 1 to 214798)
Direct Submission
Submitted (29-APR-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL
Ying, N., Song, L. and Roe, B.A.
3 (bases 1 to 214798)
Direct Submission
Submitted (14-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL
Ying, N., Song, L. and Roe, B.A.
4 (bases 1 to 214798)
Direct Submission
Submitted (15-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL
Ying, N., Song, L. and Roe, B.A.
5 (bases 1 to 214798)
Direct Submission
Submitted (27-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Nov 27, 2002 this sequence version replaced gi:24961369.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

Location/Qualifiers
1. 214798
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="10"
/clone="rp23-199g2"
/clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN
Query Match 90.6%; Score 15.4; DB 6; Length 214798;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;


```

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGAG 17
   |||||
Db 50950 GCCACGCTTGGCAAG 50966

RESULT 91
AC134138
LOCUS
DEFINITION
AC134138 216540 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-257F9, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
AC134138
VERSION
AC134138.2 GI:25007343
HTG; HTGS PHASE1; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 216540)
Muzny,D,Marle., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Ayalabechechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,K., Hampton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovat,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,L., London,P., Longacre,S., Lopez,J.,
Lorenshuwa,L., Loulsegod,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidas,M., Murphy,M., Nair,D.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokenleh,O., Okwuonu,G., Olarunpusagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
Williams,G., Willson,R., Warren,R., Warren,R., Wei,X., White,F.,
Wang,O., Wang,S., Warren,J., Wlezyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 216540)

```

```

AUTHORS
TITLE
JOURNAL
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (23-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 216540)
REFERENCE
AUTHORS
TITLE
JOURNAL
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23269344.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBWU
Center clone name: CH230-257F9
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 196236 bases at least Q40
Consensus quality: 199589 bases at least Q30
Consensus quality: 201566 bases at least Q20
Estimated insert size: 203201; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 212460: contig of 212460 bp in length
* 212461 212560: gap of unknown length
* 212561 213819: contig of 1259 bp in length
* 213820 213919: gap of unknown length
* 213920 216540: contig of 2621 bp in length.
FEATURES
source
Location/Qualifiers
1. 216540
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-257F9"
1. 1135
/note="wgs_end_extension
clone_end:T7"
5118. 5953
/note="clone_boundary
clone_end:T7
site:
end sequence:B2141658"
130387. 132038
/note="wgs_contig"
208312. 209023
/note="clone_boundary

```

```
clone_end:Sp6
site:
end_sequence: B2141659"
CR974487 .212560
/estimated_length=unknown
213820 .213919
/estimated_length=unknown

ORIGIN

Query Match          90.6%; Score 15.4; DB 12; Length 216540;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
    |||
Db 167143 GCCCAGCTTGCCGAG 167159

RESULT 92
CR974487
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-136M7 on chromosome 12, complete
sequence.
ACCESSION CR974487
VERSION CR974487.10 GI:82465399
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 218303)
AUTHORS Hunter, G.
TITLE Direct Submission
JOURNAL
COMMENT Submitted (15-NOV-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 15, 2005 this sequence version replaced gi:78271849.
-----Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vega@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
RP23-136M7 is from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
Location/Qualifiers
1. 218303
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="12"
/clone="RP23-136M7"
/clone_lib="RPCI-23"

ORIGIN

Query Match          90.6%; Score 15.4; DB 6; Length 218303;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
    |||
Db 218080 GCCCAGCTTGCCGAG 218096

RESULT 93
AC150072
LOCUS
DEFINITION Gallus gallus clone CH261-91K4, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC150072
VERSION AC150072.1 GI:49533667
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ACTIVEFIN.
SOURCE Gallus gallus (chicken)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 230993)
AUTHORS DOE Joint Genome Institute.
TITLE Unpublished
JOURNAL
COMMENT DOE Joint Genome Institute.
DOE Joint Genome Institute.
Direct Submission
Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
94598-1638, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 2965963
Center clone name: JF2-91K4
-----
Summary Statistics
Consensus quality: 227292 bases at least Q40
Consensus quality: 228380 bases at least Q30
Consensus quality: 229249 bases at least Q20
Estimated insert size: 218000; agarose-fp estimation
Estimated insert size: 230093; sum-of-contigs estimation
Quality coverage: 11.13 in Q20 bases; agarose-fp estimation
Quality coverage: 10.54 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 49795: contig of 49795 bp in length
49796 49895: gap of unknown length
49896 104846: contig of 54951 bp in length
104847 120130: contig of 15184 bp in length
120131 129521: contig of 9291 bp in length
129522 129621: gap of unknown length
129622 154193: contig of 24572 bp in length
154194 154293: gap of unknown length
154294 157923: contig of 3630 bp in length
157924 158023: gap of unknown length
158024 160903: contig of 2880 bp in length
160904 161003: gap of unknown length
```

```
* 161004 173646: contig of 12643 bp in length
* 173647 173746: gap of unknown length
* 173747 190902: contig of 17156 bp in length
* 190903 191002: gap of unknown length
* 191003 230993: contig of 39991 bp in length.
```

FEATURES
source

```
Location/Qualifiers
1..230993
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/clone="CH261-91K4"
/clone_lib="CHORI-261 Chicken BAC Library"
49796..49895
/estimated_length=unknown
104847..104946
/estimated_length=unknown
120131..120230
/estimated_length=unknown
129522..129621
/estimated_length=unknown
154194..154293
/estimated_length=unknown
157924..158023
/estimated_length=unknown
160904..161003
/estimated_length=unknown
173647..173746
/estimated_length=unknown
190903..191002
/estimated_length=unknown
```

ORIGIN

```
Query Match 90.6%; Score 15.4; DB 12; Length 230993;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 GCCCAGCCTGGCCGAG 17
|||||

Db 135912 GCCCAGCCTGGCCGAG 135896

RESULT 94

AC150066 233345 bp DNA linear HTG 01-JUL-2004
LOCUS
DEFINITION Gallus gallus clone CH261-67M3, WORKING DRAFT SEQUENCE, 7 ordered
pieces.

ACCESSION AC150066
VERSION AC150066.1 GI:49533661
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Gallus gallus (Chicken)

ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus

1 (bases 1 to 233345)
DOE Joint Genome Institute.

Unpublished
JOURNAL Unpublished

2 (bases 1 to 233345)
DOE Joint Genome Institute.

Direct Submission
Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA

94598-1698, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

COMMENT

Project Information
Center Project Name: 2956794
Center clone name: JF2-67M3

Summary Statistics
Consensus quality: 230345 bases at least Q40
Consensus quality: 231164 bases at least Q30
Consensus quality: 231877 bases at least Q20
Estimated insert size: 248000; agarose-fp estimation
Estimated insert size: 232745; sum-of-contigs estimation
Quality coverage: 12.16 in Q20 bases; agarose-fp estimation
Quality coverage: 12.96 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.

This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.

1 159426: contig of 159426 bp in length
159427 159526: gap of unknown length
159527 162471: contig of 2945 bp in length
162472 162571: gap of unknown length
162572 166152: contig of 3581 bp in length
166153 166252: gap of unknown length
166253 183843: contig of 17591 bp in length
183844 183943: gap of unknown length
183944 182052: contig of 8109 bp in length
182053 192152: gap of unknown length
192153 217751: contig of 25599 bp in length
217752 217851: gap of unknown length
217852 233345: contig of 15494 bp in length.

FEATURES
Location/Qualifiers

source

1..233345
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/clone="CH261-67M3"
/clone_lib="CHORI-261 Chicken BAC Library"
159427..159526
/estimated_length=unknown
162472..162571
/estimated_length=unknown
166153..166252
/estimated_length=unknown
183844..183943
/estimated_length=unknown
192053..192152
/estimated_length=unknown
217752..217851
/estimated_length=unknown

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 233345;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCCTGGCCGAG 17
|||||

Db 185142 GCCCAGCCTGGCCGAG 185158

RESULT 95

AC130741

LOCUS
DEFINITION Rattus norvegicus clone CH230-84A14, *** SEQUENCING IN PROGRESS
AC130741 234079 bp DNA linear HTG 10-MAY-2003
***, 4 unordered pieces.

ACCESSION

AC130741.4 GI:30522016

VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 234079)
 Muzny, D. Marie, Metkay, M. Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorjis, E., Geier, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuwa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahidatne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okwuonu, G., Olarpunsaogon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Qulroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, J., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajda, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.
 Direct Submission
 Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 234079)
 Rat Genome Sequencing Consortium.
 Submitted (14-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 234079)
 Rat Genome Sequencing Consortium.
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:25009391.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GGBE
 Center clone name: CH230-84A14
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 208845 bases at least Q40
 Consensus quality: 213470 bases at least Q30
 Consensus quality: 216683 bases at least Q20
 Estimated insert size: 217164; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 90552: contig of 90552 bp in length
 * 90553 90652: gap of unknown length
 * 90653 225351: contig of 134699 bp in length
 * 225352 225451: gap of unknown length
 * 225452 230188: contig of 4737 bp in length
 * 230189 230288: gap of unknown length
 * 230289 234079: contig of 3791 bp in length.

FEATURES
source

1. 234079
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clones="CH230-84A14"

misc_feature

1. 1516
 /note="wgs contig"

gap

90553..90652
 /estimated_length=unknown

misc_feature

148829..150032
 /note="wgs contig"

misc_feature

219786..220971
 /note="wgs contig"

gap

225352..225451
 /estimated_length=unknown

gap

230189..230288
 /estimated_length=unknown

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 234079;
 Best Local Similarity 94.1%; Pred. No. 2.4e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCCGAG 17

Db 110681 GCCCATGCTTGGCCGAG 110697

RESULT 96

AC165587

LOCUS

AC165587

DEFINITION

Bos taurus clone CH240-169117, *** SEQUENCING IN PROGRESS ***

ACCESSION

AC165587

VERSION

AC165587.2 GI:85664371

236105 bp DNA linear HTG 23-JAN-2006

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Bos taurus (cattle)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 236105)
 REFERENCE
 AUTHORS Muzny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P.,
 Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,R.,
 Beraducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
 Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
 Chacko,J., Chahrouh,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
 Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
 Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
 Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinin,H.,
 Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
 Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
 Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
 Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
 Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
 Gonzalez-Garay,M., Guevara,W., Holder,M., Hailand,W.,
 Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
 Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
 Hines,S., Hitchens,M., Hodgson,A., Hognes,M., Hollins,B.,
 Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
 Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalatus,K., Kelly,S.,
 Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
 Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
 Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
 London,P., Lopez,J., Lorensuhtewa,L., Lozado,R., Luk,T., Madu,R.,
 Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
 McClelland,H., McPherson,J., Mercadado,C., Metzker,M.,
 Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M.,
 Murray,D., Nazareth,L., Ngo,D., Nguyen,N., Norwig-Eaustagh,E.,
 Nott,A., Nwaekemele,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,
 Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
 Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
 Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J.,
 Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
 Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
 Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
 Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
 Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
 Trejos,Z., Usmani,K., Vargo,C., Verduzco,D., Villasana,D., Virk,D.,
 Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
 Wei,X., Wheeler,D., Williams,G., Williams,K., Worley,K., Wright,R.,
 Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
 Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 236105)
 REFERENCE
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 236105)
 REFERENCE
 AUTHORS .
 CONSRM Bovine Genome Sequencing Consortium
 TITLE Direct Submission
 JOURNAL Submitted (23-JAN-2006) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FJWO
 Center clone name: CH240-169117
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 211571 bases at least Q40
 Consensus quality: 213978 bases at least Q30
 Consensus quality: 216024 bases at least Q20
 Estimated insert size: 216034; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 30 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 5219: contig of 5219 bp in length
 * 5269: gap of 50 bp
 * 5270: contig of 3405 bp in length
 * 8674: gap of unknown length
 * 8775: contig of 1439 bp in length
 * 10213: contig of 50 bp
 * 10263: gap of 50 bp
 * 17437: contig of 1714 bp in length
 * 17487: gap of 50 bp
 * 70940: contig of 53453 bp in length
 * 70990: gap of 50 bp
 * 76915: contig of 5925 bp in length
 * 76965: gap of 50 bp
 * 78111: contig of 1446 bp in length
 * 78511: gap of unknown length
 * 80032: contig of 1521 bp in length
 * 81311: gap of 1279 bp
 * 93774: contig of 18463 bp in length
 * 99775: 100080: gap of 306 bp
 * 100081: contig of 1 bp in length
 * 100158: gap of 77 bp
 * 100159: contig of 1 bp in length
 * 100209: gap of 50 bp
 * 100210: contig of 4133 bp in length
 * 104342: gap of 50 bp
 * 104393: contig of 5277 bp in length
 * 104392: gap of 50 bp
 * 109670: 109669: contig of 5277 bp in length
 * 109719: gap of 50 bp
 * 109720: contig of 4990 bp in length
 * 114710: 114709: gap of 50 bp
 * 114759: contig of 3073 bp in length
 * 117832: contig of 2359 bp
 * 120191: gap of 9937 bp in length
 * 130128: contig of 50 bp
 * 130178: gap of 50 bp
 * 130179: contig of 2197 bp in length
 * 132375: gap of 7103 bp
 * 132376: 139478: contig of 9385 bp in length
 * 148863: contig of 206 bp
 * 149069: gap of 43016 bp in length
 * 192085: contig of 25591 bp in length
 * 192185: gap of unknown length
 * 192776: contig of 635 bp
 * 218411: gap of 50 bp
 * 218412: 218411: gap of 50 bp
 * 219756: 219755: contig of 1344 bp in length
 * 219805: gap of 50 bp
 * 219806: 220842: contig of 1037 bp in length
 * 220843: 220942: gap of unknown length

* 220943 223820: contig of 2878 bp in length
 * 223821 224914: gap of 1094 bp
 * 224915 226691: contig of 1777 bp in length
 * 226692 226741: gap of 50 bp
 * 226742 227763: contig of 1022 bp in length
 * 227764 227863: gap of unknown length
 * 227864 228875: contig of 1012 bp in length
 * 228876 228975: gap of unknown length
 * 228976 230395: contig of 1420 bp in length
 * 230396 230495: gap of unknown length
 * 230496 232091: contig of 1596 bp in length
 * 232092 232191: gap of unknown length
 * 232192 233466: contig of 1275 bp in length
 * 233467 233566: gap of unknown length
 * 233567 236105: contig of 2539 bp in length.

FEATURES

source

Location/Qualifiers
 1. 236105
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="CH240-169117"
 5220..5269
 /estimated_length=50
 8675..8774
 /estimated_length=unknown
 10214..10263
 /estimated_length=50
 17438..17487
 /estimated_length=50
 70941..70990
 /estimated_length=50
 76916..76965
 /estimated_length=50
 78412..78511
 /estimated_length=unknown
 80033..81311
 /estimated_length=1279
 99775..100080
 /estimated_length=306
 100082..100158
 /estimated_length=77
 100160..100209
 /estimated_length=50
 104343..104392
 /estimated_length=50
 109670..109719
 /estimated_length=50
 114710..114759
 /estimated_length=50

Query Match 90.6%; Score 15.4; DB 12; Length 236105;

Best Local Similarity 94.1%; Pred. No. 2.4e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

Db 21115 GCCCAGCTTGGCCGAG 21131

RESULT 97

AC136419/c

LOCUS AC136419 240430 bp DNA linear HTG 09-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-703, *** SEQUENCING IN PROGRESS ***.

ACCESSION

AC136419.3 GI:30467518

VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 240430)

REFERENCE

Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebirt, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensu, L., Loulsegard, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minje, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plummer, F., Poinexter, A., Popovic, D., Primus, E., Pu, L. L.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 240430)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (01-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 240430)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 9, 2003 this sequence version replaced gi:25188830.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCNS
Center clone name: CH230-703
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 227364 bases at least Q40
Consensus quality: 223046 bases at least Q30
Consensus quality: 230029 bases at least Q20
Estimated insert size: 235940: sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 240430: contig of 240430 bp in length.
  Location/Qualifiers
    1..240430
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-703"
    1..1676
      /note="wgs contig"
    184225..186084
      /note="wgs contig"
    237584..240430
      /note="wgs contig"

FEATURES             source
-----
misc_feature          1..1676
misc_feature          184225..186084
misc_feature          237584..240430

ORIGIN
Query Match          90.6%; Score 15.4; DB 12; Length 240430;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
    ||||| |||||
Db 238589 GCCACGCTTGCCGAG 238573

RESULT 98
AC106342
LOCUS              AC106342
DEFINITION         Rattus norvegicus clone CH230-154L23, WORKING DRAFT SEQUENCE, 4
unordered pieces.
ACCESSION          AC106342
VERSION            AC106342.4 GI:30581327
KEYWORDS           HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE             Rattus norvegicus (Norway rat)
ORGANISM           Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayoyi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Bunay,C., Burch,P., Burrell,K., Calderon,E.,

```

```

Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Cartroll,L., De Anda,C., Dederich,D.,
Delgado O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabis,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haviak,P., Haves,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jacksom,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karkathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuewa,L., Loulseg,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokeme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojars,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,I., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 241836)
Worley,K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 241836)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23111185.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

```


Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Center project name: GKSJ

Center clone name: CH230-154123

Assembly program: Atlas 3.0;

Consensus quality: 223364 bases at least Q40

Consensus quality: 225949 bases at least Q30

Consensus quality: 227952 bases at least Q20

Estimated insert size: 23411; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

NOTE: This is a 'working draft' sequence. It currently

consists of 4 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 156489: contig of 156489 bp in length

156490 156589: gap of unknown length

156590 238919: contig of 82230 bp in length

238920 238919: gap of unknown length

238920 240232: contig of 1313 bp in length

240233 240332: gap of unknown length

240333 241836: contig of 1504 bp in length.

Location/Qualifiers

1. 241836

/organism="Rattus norvegicus"

/db_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-154123"

1. 1042

/note="wgs contig"

156490. 156589

/estimated_length=unknown

238920. 238919

/estimated_length=unknown

240233. 240332

/estimated_length=unknown

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 241836;

Best Local Similarity 94.1%; Pred. NO. 2.4e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCGAG 17

Db 34630 GCCCATGCTTGGCGAG 34646

RESULT 99

AC111284/c

DEFINITION Rattus norvegicus clone CH230-5415, *** SEQUENCING IN PROGRESS ***,

2 unordered pieces.

AC111284

AC111284.3 GI:23264096

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 245117)

Musny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alabrooks,S., Amin,A., Anguiano,D.,

Anylebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,T., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gregorogis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kwis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuewa,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwakoelameh,O., Okwuonu,G., Olarunpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankech,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,J.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shattman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneid,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umami,K.,
Valas,R., Vera,V., Villanana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,K., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 245117)

Worley,K.C.

Direct Submission

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245117)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (21-SEP-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Sep 21, 2002 this sequence version replaced gi:21735811.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the

sequence may extend beyond the ends of the clone and there may be

contigs that consist entirely of whole genome shotgun sequence

reads. Both end sequences and whole genome shotgun sequence only

contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

```

Center project name: GLYH
Center clone name: CH230-54L5
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329
Consensus quality: 215111 bases at least Q40
Consensus quality: 218615 bases at least Q30
Consensus quality: 221156 bases at least Q20
Estimated insert size: 236907; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 243787: contig of 243787 bp in length
* 243788 243887: gap of unknown length
* 243888 245117: contig of 1230 bp in length.
FEATURES             Location/Qualifiers
     source            1..245117
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-54L5"
     misc_feature      1..1075
                     /note="wgs_end_extension"
                     clone_end:T7"
     misc_feature      1922..3168
                     /note="wgs_end_extension"
                     clone_end:T7"
     misc_feature      6356..8142
                     /note="wgs_end_extension"
                     clone_end:T7"
     misc_feature      8227..9080
                     /note="clone_boundary"
                     clone_end:T7"
                     site:ECORI
     misc_feature      end_sequence:BH350328"
                     239598..240208
                     /note="clone_boundary"
                     clone_end:Sp6
                     site:ECORI
     misc_feature      end_sequence:BH350329"
                     241850..243787
                     /note="wgs_end_extension"
                     clone_end:Sp6"
     gap              243788..243887
                     /estimated_length=unknown
ORIGIN
Query Match          90.6%; Score 15.4; DB 12; Length 245117;
Best Local Similarity 94.1%; Pred. NO. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
   |||||
Db 42590 GCCCAGCGTTGCCCA 42574

RESULT 100
AC115439/C
LOCUS          AC115439          247212 bp      DNA      linear      HTG 08-OCT-2002
DEFINITION    Rattus norvegicus clone CH230-184F24, *** SEQUENCING IN PROGRESS
              ***, 2 unordered pieces.
ACCESSION     AC115439
VERSION       AC115439.4 GI:23195508
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE        Rattus norvegicus (NorWay rat)

```

ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE
 1 (Bases 1 to 247212)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Aquiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, B.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, F., Garza, W.,
 Gbregedregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovat, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Loreschew, L., Loulseg, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okwunu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sittler, C.D., Sosa, J.,
 Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (Bases 1 to 247212)
 Worley, K.C.
 Direct Submission
 Submitted (19-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 Rat Genome Sequencing Consortium.
 3 (Bases 1 to 247212)
 Direct Submission
 Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 19, 2002 this sequence version replaced gi:21737862.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPWF
Center clone name: CH230-184F24
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 218155 bases at least Q40
Consensus quality: 220835 bases at least Q30
Consensus quality: 222648 bases at least Q20
Estimated insert size: 238245; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 115755: contig of 115755 bp in length
* 115756 115855: gap of unknown length
* 115856 247212: contig of 131357 bp in length.

FEATURES

source
1..247212
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-184F24"
1..1360
/note="wgs contig"
114369..115755
/note="wgs contig"
115756..115855
/estimated_length=unknown
200330..201623
/note="wgs contig"

ORIGIN

Query Match 90.6%; Score 15.4; DB 12; Length 247212;
Best Local Similarity 94.1%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
|||||
Db 144996 GCCCAGCGTGGCCGAG 144980

Search completed: June 10, 2006, 18:20:43
Job time : 518.533 secs

THIS PAGE IS BLANK

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:16:40 ; Search time 126.367 seconds
(without alignments)
937.971 Million cell updates/sec

Title: US-10-600-816-31

Perfect score: 17

Sequence: 1 gccacgctggccgag 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

N Geneseq_8:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2002bs:*

6: Geneseqn2002as:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2003ds:*

11: Geneseqn2004as:*

12: Geneseqn2004bs:*

13: Geneseqn2004cs:*

14: Geneseqn2005s:*

15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	100.0	17	12	ADI28528 Human GPC
2	17	100.0	40	12	ADI28527 Human GPC
3	17	100.0	497	11	ADT96418 Colon can
4	17	100.0	497	11	ADT96418 Human GPC
5	17	100.0	552	11	ADT96292 Colon can
6	17	100.0	552	11	ADT96292 Human GPC
7	17	100.0	603	3	AZ90046 Hydrophob
8	17	100.0	620	6	ABQ58527 Human col
9	17	100.0	634	6	ABQ58527 Human col
10	17	100.0	642	10	ADT22328 Breast ca
11	17	100.0	1074	12	ADT22328 Human GPC
12	17	100.0	1212	3	AZ61776 cDNA enco
13	17	100.0	1212	4	AAC99709 Skin cell
14	17	100.0	1212	6	ABL34861 Human GPC
15	17	100.0	1228	8	ACA04775 cDNA enco
16	17	100.0	1460	14	AEA00112 Human TAT
17	17	100.0	1460	14	AEA00632 Human TAT
18	17	100.0	1619	4	AAF58615 Human REC

19	17	100.0	1718	3	AAZ90056	Aaz90056 Hydrophob
20	17	100.0	1788	10	ADF70573	Acf70573 Orphan re
21	17	100.0	2302	6	ABT10173	Abt10173 Human bre
22	17	100.0	2302	8	ACC58386	Acc58386 Human GPC
23	17	100.0	2302	10	ADD93240	Ad93240 RAIG1 cod
24	17	100.0	2302	10	ADL24773	Adl24773 Intestina
25	17	100.0	2316	10	ABT31923	Abt31923 Human bre
26	17	100.0	2446	4	AAH14688	Aah14688 Human GPC
27	17	100.0	2456	8	ABZ42832	Abz42832 Human G p
28	17	100.0	2456	10	ACC72695	Acc72695 Human can
29	17	100.0	2456	11	ADN39301	Adn39301 Cancer/an
30	17	100.0	2456	12	ADI28525	Adi28525 Human GPC
31	17	100.0	2456	12	ADI28459	Adi28459 Human GPC
32	17	100.0	2456	13	ADR48221	Adr48221 Human ret
33	17	100.0	2456	13	ACN39388	Acn39388 Tumour-as
34	17	100.0	2456	13	ADR43992	Adr43992 Human bre
35	17	100.0	2456	13	ADU06126	Adu06126 Novel bro
36	17	100.0	2456	14	AEA00088	Aea00088 Human TAT
37	17	100.0	2456	14	AEA00608	Aea00608 Human TAT
38	17	100.0	2456	14	AED47480	Aed47480 Retinoic
39	17	100.0	2593	6	ABQ54954	Abq54954 Human ova
40	17	100.0	3371	4	AAH72766	Aah72766 Human cer
41	17	100.0	4239	11	ACN89274	Acn89274 Breast ca
42	17	100.0	6730	13	ADX97494	Adx97494 Pancreati
43	15.4	90.6	1527	11	ABD12717	Abd12717 Pseudomon
44	15.4	90.6	2306	6	ABQ70591	Abq70591 Listeria
45	15.4	90.6	2976	14	AEA48186	Aea48186 DNA Of si
46	15.4	90.6	4183	2	AXH81946	Aax81946 DNA encod
47	15.4	90.6	4222	12	ADQ24610	Adq24610 Human sof
48	15.4	90.6	4232	13	ADQ7536	Adq7536 Full leng
49	15.4	90.6	9699	5	ABA21144	Ab21144 Human ner
50	15.4	90.6	17457	5	ABA21143	Ab21143 Human ner
51	15.4	90.6	17458	5	ABA21142	Ab21142 Human ner
52	15.4	90.6	63761	13	ABD33426	Abd33426 Murine ca
53	15.4	90.6	68732	13	ABD33428	Abd33428 Human can
54	15.4	90.6	95269	6	ABQ67195	Abq67195 4
55	15.4	90.6	110000	6	ABQ69245	Abq69245 13
56	15.4	90.6	110000	6	ABQ67195	Abq67195 3
57	15.4	90.6	110000	6	ABA03041	Ab3041 13
58	15.4	90.6	349980	6	ABQ81846	Abq81846 Bifidobac
59	15.4	90.6	349980	6	ABQ81847	Abq81847 Bifidobac
60	15	88.2	337	5	ABV57592	Abv57592 Human pro
61	15	88.2	3303	13	ADT19568	Adt19568 Plant CDN
62	14.4	84.7	344	4	AAS37371	Aas37371 Novel hum
63	14.4	84.7	546	4	AAH09788	Aah09788 Human GPC
64	14.4	84.7	570	5	ABV54022	Abv54022 Human pro
65	14.4	84.7	680	4	AAH08354	Aah08354 Human GPC
66	14.4	84.7	992	8	AAH5183	Aal5183 Verticill
67	14.4	84.7	1520	2	AAV63901	Aav63901 Helicobac
68	14.4	84.7	1531	8	ACD05884	Ac05884 Novel hum
69	14.4	84.7	1675	2	AAV63900	Aav63900 Helicobac
70	14.4	84.7	1889	6	ABZ35358	Abz35358 Human gen
71	14.4	84.7	3681	13	ACN40151	Acn40151 Tumour-as
72	14.4	84.7	48169	14	AED48654	Aed48654 Streptomy
73	14.4	84.7	49312	3	AAH51594	Aah51594 Human gen
74	14.4	84.7	89378	12	ADN46845	Adn46845 20
75	14.4	84.7	89378	12	ADN46123	Adn46123 20
76	14.4	84.7	89378	12	ADN46123	Adn46123 20
77	14.4	84.7	89378	12	ADN46123	Adn46123 20
78	14.4	84.7	110000	12	ADN47591	Adn47591 Thermococ
79	14.4	84.7	110000	12	ADN47209	Adn47209 Thermococ
80	14.4	84.7	110000	12	ADN47209	Adn47209 Thermococ
81	14.4	84.7	32602	13	ABD12843	Abd12843 Human can
82	14	82.4	338	6	ABN23393	Abn23393 Human ORF
83	14	82.4	458	9	ACH41474	Ach41474 Human foe
84	14	82.4	508	12	ADN12918	Adn12918 Human pro
85	14	82.4	706	12	ADN13857	Adn13857 Human pro
86	14	82.4	1234	3	AAC47788	Aac47788 Arabidops
87	14	82.4	110000	12	ADN46845	Adn46845 16
88	14	82.4	110000	12	ADN47591	Adn47591 04
89	14	82.4	110000	12	ADN46123	Adn46123 16
90	14	82.4	110000	12	ADN47209	Adn47209 04
91	14	82.4	110000	12	ADN46464	Adn46464 16

92	14	82.4	110000	12	ADN47960_04	Continuation (5 of	13.8	81.2	1578	2	AAQ61817	Aaq61817 Open read
93	14	82.4	349980	6	ABQ81848	Abq81848 Bifidobac	13.8	81.2	1599	14	AEA61255	Aea61255 Human CA2
94	14	82.4	349980	6	ABQ81844	Abq81844 Bifidobac	13.8	81.2	1608	11	ABD01965	Abd01965 Pseudomon
95	13.8	81.2	47	3	AAZ68079	Aaz68079 Human map	13.8	81.2	1616	3	AAC81948	Aac81948 Backbone
96	13.8	81.2	228	6	ABN78126	Abn78126 Human gly	13.8	81.2	1632	14	AEA35610	Aea35610 Pseudomon
97	13.8	81.2	247	10	ADA48755	Ada48755 Human acy	13.8	81.2	1635	12	ADP53526	Adp53526 Human hex
98	13.8	81.2	283	10	ABX84111	Abx84111 Corn ear-	13.8	81.2	1635	12	ADF53519	Adf53519 Human hex
99	13.8	81.2	290	12	ADP95237	Adp95237 Cotton ex	13.8	81.2	1635	14	AEC36861	Aec36861 Human hex
100	13.8	81.2	293	10	ABX86649	Abx86649 Corn ear-	13.8	81.2	1635	14	AEC36868	Aec36868 Human hex
101	13.8	81.2	350	6	ABL78156	Ab178156 Human ova	13.8	81.2	1653	11	ABD02507	Abd02507 Pseudomon
102	13.8	81.2	363	6	ABN75772	Abn75772 Human ORF	13.8	81.2	1668	14	AEA35622	Aea35622 P. aerugi
103	13.8	81.2	378	10	ABZ38151	Abz38151 N. gonorr	13.8	81.2	1671	14	AED73829	Aed73829 Human pla
104	13.8	81.2	385	2	AAT49143	Aat49143 Partial D	13.8	81.2	1672	14	AEA23969	Aea23969 Human PRO
105	13.8	81.2	391	9	ACH17606	Ach17606 Human adu	13.8	81.2	1680	4	AAF61077	Aaf61077 P. putida
106	13.8	81.2	417	4	ABL23399	Ab123399 Drosophil	13.8	81.2	1734	13	ADT48713	Adt48713 Bacterial
107	13.8	81.2	459	10	ABZ38162	Abz38162 N. gonorr	13.8	81.2	1746	14	AEA51326	Aea51326 Human gen
108	13.8	81.2	522	4	AAK82910	Aak82910 Human imm	13.8	81.2	1784	10	ADC64481	Adc64481 Human inh
109	13.8	81.2	526	9	ACH16268	Ach16268 Human adu	13.8	81.2	1807	11	ACN91739	Acn91739 Breast ca
110	13.8	81.2	543	11	ADQJ11769	Adqj11769 Rice cDNA	13.8	81.2	1837	12	ADJ66064	Adj66064 Human ion
111	13.8	81.2	546	11	ADGJ11415	Adgj11415 Rice DNA	13.8	81.2	1839	13	ADT79424	Adt79424 Human hex
112	13.8	81.2	574	9	ACH41181	Ach41181 Human foe	13.8	81.2	1857	12	ADF53531	Adf53531 Human hex
113	13.8	81.2	600	5	AAD04923	Aad04923 Pyromycin	13.8	81.2	1857	13	ACN38246	Acn38246 Tumour-as
114	13.8	81.2	600	5	AAF85069	Aaf85069 Nucleotid	13.8	81.2	1857	13	ADP23310	Adp23310 PRO polyP
115	13.8	81.2	600	6	ABK50562	Abk50562 Streptom	13.8	81.2	1857	14	ADZ26230	Adz26230 Human gen
116	13.8	81.2	600	8	AAD49162	Aad49162 Streptom	13.8	81.2	1857	14	ADZ48954	Adz48954 Insulin s
117	13.8	81.2	600	15	AEF20958	Aef20958 Mouse Pur	13.8	81.2	1857	14	AEC36873	Aec36873 Nucleic a
118	13.8	81.2	677	6	ABS51073	Abs51073 cDNA enco	13.8	81.2	1892	8	ABZ70965	Abz70965 Human cer
119	13.8	81.2	677	10	ADJ83022	Adj83022 Human NOV	13.8	81.2	1901	12	ADH35387	Adh35387 ENZM enco
120	13.8	81.2	751	6	ABQ82353	Abq82353 Human NOV	13.8	81.2	1958	3	AAD01033	Aad01033 Human cyc
121	13.8	81.2	755	10	AD48363	Ad48363 Rat gene	13.8	81.2	1972	12	ADQ86211	Adq86211 Human tum
122	13.8	81.2	795	14	ACL69001	Ac169001 M. xanthu	13.8	81.2	1972	13	ADQ85047	Adq85047 Human tum
123	13.8	81.2	799	4	AAI94230	Aai94230 Human neu	13.8	81.2	1972	13	ACN38859	Acn38859 Tumour-as
124	13.8	81.2	867	8	ACA33854	Ac433854 Prokaryot	13.8	81.2	1976	3	AAA76169	Aaa76169 Human ACA
125	13.8	81.2	873	14	AED47678	Aed47678 2, 4-dichl	13.8	81.2	1976	12	ADP13311	Adp13311 Renal cel
126	13.8	81.2	873	14	AED47678	Aed47678 2, 4-dichl	13.8	81.2	1976	12	ADS00240	Ads00240 Human dia
127	13.8	81.2	951	4	AA502361	Aa502361 DNA enco	13.8	81.2	1976	13	ADZ49273	Adz49273 Insulin s
128	13.8	81.2	966	11	ADT45923	Adt45923 Bacterial	13.8	81.2	1976	14	AED43572	Aed43572 DGAT1 (d1
129	13.8	81.2	999	13	ABD07657	Abd07657 Pseudomon	13.8	81.2	1983	11	ABD01927	Abd01927 Pseudomon
130	13.8	81.2	1027	4	AAI92566	Aai92566 Human pol	13.8	81.2	1992	10	ADG76394	Adg76394 Human Inc
131	13.8	81.2	1050	13	ADS55972	Ads55972 Bacterial	13.8	81.2	2000	8	ADA73460	Ada73460 Rice gene
132	13.8	81.2	1056	13	ADT42342	Adt42342 Bacterial	13.8	81.2	2000	11	ABD07822	Abd07822 Pseudomon
133	13.8	81.2	1101	8	ACA42781	Ac42781 Prokaryot	13.8	81.2	2049	1	AAAN80093	Aan80093 Sequence
134	13.8	81.2	1111	10	ADRE35068	Ade35068 IRES-puro	13.8	81.2	2058	2	AAQ97857	Aaq97857 P. aerugi
135	13.8	81.2	1147	13	ADR24980	Adr24980 Breast ca	13.8	81.2	2058	2	AAQ97857	Aaq97857 P. aerugi
136	13.8	81.2	1150	3	AAZ51265	Aaz51265 Human RNA	13.8	81.2	2085	2	AAQ90718	Aaq90718 Human rh3
137	13.8	81.2	1150	10	ADC32268	Adc32268 Human nov	13.8	81.2	2085	4	AAAS59253	Aas59253 Human cDN
138	13.8	81.2	1161	5	AAF30414	Aaf30414 Moraxella	13.8	81.2	2085	6	ABA90922	Ab90922 Human pol
139	13.8	81.2	1161	15	AEF63884	Aef63884 pAC-Beta	13.8	81.2	2085	14	ADW09180	Adw09180 cDNA clon
140	13.8	81.2	1168	13	ADU05974	Adu05974 Novel bro	13.8	81.2	2120	6	ABS79041	Abs79041 E. coli C
141	13.8	81.2	1168	14	ADV70023	Adv70023 Human RNA	13.8	81.2	2120	10	ADH80608	Adh80608 Escherich
142	13.8	81.2	1168	14	ADX07214	Adx07214 Cyclin-de	13.8	81.2	2128	8	ABZ36131	Abz36131 Human sec
143	13.8	81.2	1168	14	ADY56978	Ady56978 Human Rib	13.8	81.2	2129	5	ABV30263	Abv30263 Human pro
144	13.8	81.2	1168	15	AEF75087	Aef75087 Human pol	13.8	81.2	2165	12	ADP10476	Adp10476 Reference
145	13.8	81.2	1176	11	ABD01943	Abd01943 Pseudomon	13.8	81.2	2169	12	ADL12586	Adl12586 Human ste
146	13.8	81.2	1182	12	ADL02580	Adl02580 DNA enco	13.8	81.2	2172	13	ADS14790	Ads14790 Pseudomon
147	13.8	81.2	1210	14	AEQ04117	Aee04117 IRES-Puro	13.8	81.2	2174	6	AAD29645	Aad29645 Human dia
148	13.8	81.2	1234	2	AAQ13722	Aaq13722 Lycopen	13.8	81.2	2184	14	ACL68653	Ac168653 M. xanthu
149	13.8	81.2	1234	2	AAQ13723	Aaq13723 Lycopen	13.8	81.2	2207	2	AAAS58986	Aas58986 Mouse tra
150	13.8	81.2	1235	2	AAT40795	Aat40795 Lycopen	13.8	81.2	2219	13	ADS58712	Ads58712 Bacterial
151	13.8	81.2	1235	2	AAT40796	Aat40796 Recombina	13.8	81.2	2375	14	AEA35616	Aea35616 P. aerugi
152	13.8	81.2	1235	2	AAT91548	Aat91548 Genetical	13.8	81.2	2417	4	ABL23398	Ab123398 Drosophil
153	13.8	81.2	1235	2	AAT91548	Aat91548 Genetical	13.8	81.2	2417	4	ABL23398	Ab123398 Drosophil
154	13.8	81.2	1251	11	ADQ02415	Adq02415 Pseudomon	13.8	81.2	2445	10	ABA01061	Ab01061 Triple fu
155	13.8	81.2	1263	5	ABD08890	Abd08890 Mycobacte	13.8	81.2	2541	13	ADS58598	Ads58598 Bacterial
156	13.8	81.2	1267	3	AAZ59301	Aaz59301 Rat insul	13.8	81.2	2573	12	ADO64905	Adg64905 Novel hum
157	13.8	81.2	1267	12	ADM47807	Adm47807 Polynucle	13.8	81.2	2732	13	ADX54728	Adx54728 Plant ful
158	13.8	81.2	1290	8	ACA44303	Aca44303 Prokaryot	13.8	81.2	2787	4	ABL10299	Ab110299 Drosophil
159	13.8	81.2	1294	3	ADQ01037	Adq01037 Human cyc	13.8	81.2	2847	11	ABD02432	Abd02432 Pseudomon
160	13.8	81.2	1335	8	ADA69556	Ada69556 Rice gene	13.8	81.2	3133	2	AAQ90652	Aaq90652 Eph-relat
161	13.8	81.2	1384	13	ADT16994	Adt16994 Plant cDN	13.8	81.2	3179	13	ADRO6816	Adro6816 Full leng
162	13.8	81.2	1442	3	AAD01034	Aad01034 Human cyc	13.8	81.2	3360	4	ABL14365	Ab114365 Drosophil
163	13.8	81.2	1450	3	AAD01036	Aad01036 Human cyc	13.8	81.2	3453	14	ACL68250	Ac168250 M. xanthu
164	13.8	81.2	1488	8	ACA23486	Aca23486 Prokaryot	13.8	81.2	3486	8	ACA45671	Ac45671 Prokaryot

384	13.8	81.2	10325	2	AAX76516	Aax76516 Plasmid c	c 457	13.8	81.2	27435	13	ADR83408	Adr83408 Human ena
385	13.8	81.2	10325	9	ADB17612	Adb17612 Plasmid u	c 458	13.8	81.2	27642	4	AAK77486	Aak77486 Human imm
386	13.8	81.2	10330	12	ADL671154	Adl671154 Plasmid p	c 459	13.8	81.2	28058	14	ACL64789	ACL64789 M. xanthu
387	13.8	81.2	10477	12	ADL671152	Adl671152 Plasmid p	c 460	13.8	81.2	28690	5	AAF57718	Aaf57718 Human IL4
388	13.8	81.2	10516	12	ADL671150	Adl671150 Plasmid p	c 461	13.8	81.2	28762	14	ACL64769	ACL64769 M. xanthu
389	13.8	81.2	10553	8	AAD49160	Aad49160 CET710 ve	c 462	13.8	81.2	29899	14	ACL64802	ACL64802 M. xanthu
390	13.8	81.2	10554	14	ADW171849	Adw171849 APEX-3p G	c 463	13.8	81.2	32421	15	AEG11312	Aeg11312 Mycobacte
391	13.8	81.2	10561	12	ADL671148	Adl671148 Plasmid p	c 464	13.8	81.2	32846	10	ADG86868	Adg86868 Human GPC
392	13.8	81.2	10615	12	ADL671175	Adl671175 Plasmid p	c 465	13.8	81.2	36181	10	ADB74371	Adb74371 Mycobacte
393	13.8	81.2	10637	13	ADS34424	Ads34424 POSH prot	c 466	13.8	81.2	39707	13	ADJ39405	Adj39405 Herpes s1
394	13.8	81.2	10741	14	AED41937	Aed41937 Human spe	c 467	13.8	81.2	40558	12	ADJ47542	Adj47542 Human H19
395	13.8	81.2	10742	13	ACN38544	Acn38544 Tumour-as	c 468	13.8	81.2	40772	13	ADM39403	Adm39403 Herpes s1
396	13.8	81.2	10742	13	ADS34416	Ads34416 POSH prot	c 469	13.8	81.2	54779	13	ADV15286	Adv15286 Human ost
397	13.8	81.2	10742	14	ADY20140	Ady20140 DNA encod	c 470	13.8	81.2	54780	13	ADV15298	Adv15298 Human ost
398	13.8	81.2	10771	12	ADL671153	Adl671153 Plasmid p	c 471	13.8	81.2	55196	14	ADZ13001	Adz13001 Human can
399	13.8	81.2	10921	12	ADL671151	Adl671151 Plasmid p	c 472	13.8	81.2	60815	11	ACN43882	Acn43882 Human gen
400	13.8	81.2	10961	12	ADL671149	Adl671149 Plasmid p	c 473	13.8	81.2	63115	10	ADC85419	Adc85419 Mouse Tle
401	13.8	81.2	11005	12	ADL671147	Adl671147 Plasmid p	c 474	13.8	81.2	63294	9	ADA02939	Ada02939 Mouse Tle
402	13.8	81.2	11059	12	ADL671176	Adl671176 Plasmid p	c 475	13.8	81.2	63294	10	ADB72677	Adb72677 Mouse Tle
403	13.8	81.2	11105	8	AAD54698	Aad54698 pBDUpuro1	c 476	13.8	81.2	63294	12	ADM74534	Adm74534 Murine ca
404	13.8	81.2	11110	12	ADF53535	Adf53535 HIV Hex v	c 477	13.8	81.2	66566	3	AAA53450	Aaa53450 Human th1
405	13.8	81.2	11110	14	AEC36877	Aec36877 Nucleic a	c 478	13.8	81.2	70419	8	AAD56111	Aad56111 Human WNT
406	13.8	81.2	11160	8	AAD54674	Aad54674 pBDUpuro8	c 479	13.8	81.2	70419	9	ADA02473	Ada02473 Human WNT
407	13.8	81.2	11160	10	ABT16401	Abt16401 Recombina	c 480	13.8	81.2	70419	11	ADB72212	Adb72212 Human WNT
408	13.8	81.2	11189	14	ADW171853	Adw171853 APEX-3Pmo	c 481	13.8	81.2	70779	11	ACN44522	Acn44522 Human gen
409	13.8	81.2	11234	15	AED40411	Aed40411 PRRSV rel	c 482	13.8	81.2	73882	13	ADS73531	Ads73531 tcp gene
410	13.8	81.2	11646	8	AAD49170	Aad49170 CET1020 v	c 483	13.8	81.2	74407	14	ADL12490	Adl12490 Human can
411	13.8	81.2	11705	2	AAV621160	Aav621160 HSV-2 str	c 484	13.8	81.2	76146	10	ADL13505	Adl13505 Osteoarth
412	13.8	81.2	11916	12	ADO06861	Ado06861 Virucidal	c 485	13.8	81.2	80420	13	ABD33096	Abd33096 Human can
413	13.8	81.2	11941	8	AAD54696	Aad54696 pBDUpuro1	c 486	13.8	81.2	100848	4	AAF28552	Aaf28552 Genomic f
414	13.8	81.2	11998	8	AAD54694	Aad54694 pBDUpuro4	c 487	13.8	81.2	109147	12	ADQ97718	Adq97718 Mouse can
415	13.8	81.2	12088	8	AAD54670	Aad54670 pBDUpuro4	c 488	13.8	81.2	110000	4	AAI99682	AAI99682 10
416	13.8	81.2	12088	10	ABT16397	Abt16397 Recombina	c 489	13.8	81.2	110000	4	AAI99683	AAI99683 10
417	13.8	81.2	12091	8	AAD49169	Aad49169 CET1010 v	c 490	13.8	81.2	110000	11	ADM27081	Adm27081 Hyperther
418	13.8	81.2	12263	12	ADF53534	Adf53534 Human bet	c 491	13.8	81.2	115851	11	ACN43936	Acn43936 Mouse gen
419	13.8	81.2	12263	14	AEC36876	Aec36876 Nucleic a	c 492	13.8	81.2	117213	2	AAV62176	Aav62176 HSV-2 str
420	13.8	81.2	12514	12	ADO36181	Ado36181 Protein e	c 493	13.8	81.2	117213	12	ADM29048	Adm29048 Human int
421	13.8	81.2	12588	8	AAD54693	Aad54693 pBDUpuro3	c 494	13.8	81.2	130427	13	ADV35011	Adv35011 Murine CD
422	13.8	81.2	12591	8	AAD54673	Aad54673 pBDUpuro7	c 495	13.8	81.2	152261	14	AEC07415	Aec07415 DNA of HS
423	13.8	81.2	12591	10	ABT16400	Abt16400 Recombina	c 496	13.8	81.2	152261	14	AED00550	Aed00550 Herpes so
424	13.8	81.2	12680	8	AAD54669	Aad54669 pBDUpuro3	c 497	13.8	81.2	154746	6	AAD25519	Aad25519 Human her
425	13.8	81.2	12680	10	ABT16396	Abt16396 Recombina	c 498	13.8	81.2	154746	14	AEC07416	Aec07416 DNA of HS
426	13.8	81.2	12700	2	AAV62133	Aav62133 HSV-2 str	c 499	13.8	81.2	159144	10	ADL13748	Adl13748 Osteoarth
427	13.8	81.2	12753	12	ADM98599	Adm98599 Geranylge	c 500	13.8	81.2	170001	14	AED18493	Aed18493 Fibroblc
428	13.8	81.2	13068	12	ADL71890	Adl71890 Temperatu	c 501	13.8	81.2	220756	12	ADG86300	Adg86300 Human SMR
429	13.8	81.2	13198	3	AAZ36211	Aaz36211 Nucleotid	c 502	13.8	81.2	233380	11	ACN44282	Acn44282 Human gen
430	13.8	81.2	13547	8	AAD49161	Aad49161 CET720 ve	c 503	13.8	81.2	257645	12	ADQ97289	Adq97289 Human can
431	13.8	81.2	13766	6	AAD46790	Aad46790 pGRN145 p	c 504	13.8	81.2	326002	13	ABD32843	Abd32843 Human can
432	13.8	81.2	14262	8	AAD54675	Aad54675 CET720GFP	c 505	13.8	81.2	349980	6	ABQ81845	Abq81845 Bifidobac
433	13.8	81.2	14262	10	ABT16402	Abt16402 Recombina	c 506	13.4	78.8	20	6	AAD40871	Ad40871 Human hep
434	13.8	81.2	14988	14	ADY64327	Ady64327 Human ape	c 507	13.4	78.8	20	6	AAD40689	Ad40689 Human hep
435	13.8	81.2	15081	12	ADL71909	Adl71909 Expressio	c 508	13.4	78.8	130	12	ACH88494	Ach88494 Human gen
436	13.8	81.2	15083	14	AEA16305	Aea16305 DNA seque	c 509	13.4	78.8	210	8	ABX55961	Abx55961 Bovine ES
437	13.8	81.2	15110	14	AEA16307	Aea16307 DNA seque	c 510	13.4	78.8	234	2	AAT21157	Aat21157 Human gen
438	13.8	81.2	16847	12	ADO07464	Ado07464 Japanese	c 511	13.4	78.8	238	10	ADL24551	Adl24551 Intestina
439	13.8	81.2	17290	6	ABQ75117	Abq75117 Human ach	c 512	13.4	78.8	243	6	ABN25451	Abn25451 Human ORF
440	13.8	81.2	17344	4	ABL20118	Ab120118 Drosophil	c 513	13.4	78.8	277	3	AAC30908	Aac30908 Human sec
441	13.8	81.2	17358	4	ABL10306	Ab110306 Drosophil	c 514	13.4	78.8	285	6	ABN79537	Abn79537 Human iso
442	13.8	81.2	17402	13	ADS75098	Ads75098 Plasmid p	c 515	13.4	78.8	334	4	AAK60978	Aak60978 Human imm
443	13.8	81.2	17402	13	ADT92670	Adt92670 Plasmid p	c 516	13.4	78.8	341	4	AAI84396	Aai84396 Human pol
444	13.8	81.2	17402	14	AEC31852	Aec31852 Plasmid p	c 517	13.4	78.8	342	14	ACL71406	ACL71406 M. xanthu
445	13.8	81.2	17402	15	AEE32467	Aee32467 Plasmid p	c 518	13.4	78.8	347	4	AAI85485	Aai85485 Human pol
446	13.8	81.2	17753	12	ADL71910	Adl71910 Expressio	c 519	13.4	78.8	348	5	ADI74147	Adi74147 Human ova
447	13.8	81.2	18116	13	ADS75097	Ads75097 Plasmid p	c 520	13.4	78.8	348	5	ADI67769	Adi67769 Human ova
448	13.8	81.2	18116	13	ADT92669	Adt92669 Plasmid p	c 521	13.4	78.8	387	4	AAI89979	Aai89979 Human pol
449	13.8	81.2	18116	14	AEC31851	Aec31851 Plasmid p	c 522	13.4	78.8	392	2	AAI22608	Aai22608 Human gen
450	13.8	81.2	18116	15	AEE32466	Aee32466 Plasmid p	c 523	13.4	78.8	402	10	ADD67014	Add67014 Human lun
451	13.8	81.2	18476	4	AEA61199	Aea61199 Human CA2	c 524	13.4	78.8	402	10	ADE88268	Ade88268 Human lun
452	13.8	81.2	20216	4	AAS59596	Aas59596 Propionib	c 525	13.4	78.8	414	5	ADL39384	Adl39384 Human ova
453	13.8	81.2	20216	8	ACF64525	Acf64525 Propionib	c 526	13.4	78.8	433	4	AAK59246	Aak59246 Human imm
454	13.8	81.2	21293	12	ADQ97295	Adq97295 Human can	c 527	13.4	78.8	447	4	AAI91871	Aai91871 Human pol
455	13.8	81.2	25382	13	ACN37244	Acn37244 Human per	c 528	13.4	78.8	461	6	AAS97228	Aas97228 Neisseria
456	13.8	81.2	25761	6	ABQ75123	Abq75123 Human p15	c 529	13.4	78.8	467	14	ACL62386	ACL62386 Human col

C 530	13.4	78.8	485	2	AAQ26994	AaQ26994 HCV gene	C 603	13.4	78.8	1548	6	AAQ20024	AaQ20024 Bacillus
C 531	13.4	78.8	491	4	AAI90610	Abi90610 Human pol	C 604	13.4	78.8	1548	6	AAI72213	Abi72213 Bacillus
C 532	13.4	78.8	501	4	ABQ98960	AbQ98960 Human ORF	C 605	13.4	78.8	1548	6	ABL96209	AbL96209 Termamyl-
C 533	13.4	78.8	504	5	ADL42694	AdL42694 Human ova	C 606	13.4	78.8	1548	10	ADJ92081	AdJ92081 B. stearo
C 534	13.4	78.8	505	13	ACN49678	Acn49678 Cotton pr	C 607	13.4	78.8	1548	14	ADZ59161	AdZ59161 Bacillus
C 535	13.4	78.8	509	12	ACH72146	Ach72146 Human gen	C 608	13.4	78.8	1548	14	AEC78247	Aec78247 Variant a
C 536	13.4	78.8	544	12	ACH74794	Ach74794 Human gen	C 609	13.4	78.8	1549	3	AAA46361	Aaa46361 Nucleotid
C 537	13.4	78.8	563	4	AAH10207	Aah10207 Human cDN	C 610	13.4	78.8	1615	10	AAAD62370	Aad62370 Human pro
C 538	13.4	78.8	564	2	ABQ27016	AbQ27016 HK14. 1/1	C 611	13.4	78.8	1615	10	ABX16007	Abx16007 Human pro
C 539	13.4	78.8	600	6	ABQ57262	AbQ57262 Human col	C 612	13.4	78.8	1615	14	AED64259	Aed64259 Human ser
C 540	13.4	78.8	600	6	ABK78575	AbK78575 Bacillus	C 613	13.4	78.8	1626	14	AED95552	Aed95552 Coding se
C 541	13.4	78.8	619	6	ABQ57219	AbQ57219 Human col	C 614	13.4	78.8	1641	5	AAH67641	Aah67641 C glutami
C 542	13.4	78.8	626	4	AAAD03891	Aad03891 CpG islan	C 615	13.4	78.8	1641	14	ACL72454	Acl72454 M. xanthu
C 543	13.4	78.8	700	8	ACF64337	Acf64337 Human IFN	C 616	13.4	78.8	1650	8	ABX08455	Abx08455 DNA encod
C 544	13.4	78.8	759	6	ABQ55407	AbQ55407 Human ova	C 617	13.4	78.8	1650	8	ABX08461	Abx08461 DNA encod
C 545	13.4	78.8	765	4	AAK91701	Aak91701 Human cDN	C 618	13.4	78.8	1650	13	ADU07393	Adu07393 DNA seque
C 546	13.4	78.8	765	4	AAK93719	Aak93719 Human cDN	C 619	13.4	78.8	1650	13	ADU07381	Adu07381 DNA seque
C 547	13.4	78.8	765	12	ADL28128	AdL28128 5' end of	C 620	13.4	78.8	1650	14	AED95531	Aed95531 Bacillus
C 548	13.4	78.8	765	12	ADL30146	AdL30146 3' end of	C 621	13.4	78.8	1680	1	AAAN70926	Aan70926 Plasmid p
C 549	13.4	78.8	835	12	ACH90888	Ach90888 Human gen	C 622	13.4	78.8	1680	14	AEA79691	Aea79691 Cellulomo
C 550	13.4	78.8	835	12	ACH90888	Ach90888 Human gen	C 623	13.4	78.8	1700	1	AAAN70538	Aan70538 Sequence
C 551	13.4	78.8	858	2	AAV45175	Aav45175 Human LEA	C 624	13.4	78.8	1700	10	ADE79076	Ade79076 Human pro
C 552	13.4	78.8	861	11	ABD01313	Abd01313 Klebsiell	C 625	13.4	78.8	1719	1	AAAN70920	Aan70920 Plasmid p
C 553	13.4	78.8	897	2	AAQ25748	AaQ25748 Non-A, No	C 626	13.4	78.8	1719	1	AAAN70920	Aan70920 Plasmid p
C 554	13.4	78.8	897	2	AAQ25748	AaQ25748 Non-A, No	C 627	13.4	78.8	1746	4	AAK70288	Aak70288 Human imm
C 555	13.4	78.8	915	13	ADX36206	Adx36206 Plant ful	C 628	13.4	78.8	1746	4	AAK70289	Aak70289 Human imm
C 556	13.4	78.8	921	13	ADT04075	Adt04075 Human pro	C 629	13.4	78.8	1764	14	AED46904	Aed46904 Membrane
C 557	13.4	78.8	936	6	ABZ82525	Abz82525 Human sec	C 630	13.4	78.8	1769	12	ADJ57605	Adj57605 Human hep
C 558	13.4	78.8	936	13	ADX33318	Adx33318 Plant ful	C 631	13.4	78.8	1783	4	RAD13167	Rad13167 Human hep
C 559	13.4	78.8	943	14	ABEB35259	Aeb35259 Human Gef	C 632	13.4	78.8	1783	5	AAAS43103	Aas43103 CDNA enc
C 560	13.4	78.8	974	13	ADX51739	Adx51739 Plant ful	C 633	13.4	78.8	1783	6	ABQ79532	AbQ79532 Human hep
C 561	13.4	78.8	984	13	ADX47107	Adx47107 Plant ful	C 634	13.4	78.8	1783	9	ADAI17045	Adai17045 CDNA enc
C 562	13.4	78.8	1038	8	AAH99136	Aah99136 Human EST	C 635	13.4	78.8	1783	10	ADB75353	Adb75353 Prostata
C 563	13.4	78.8	1176	8	ACA18988	Aca18988 Prokaryot	C 636	13.4	78.8	1783	10	ADD18429	Add18429 Human pro
C 564	13.4	78.8	1176	13	ADU67836	Adu67836 DNA encod	C 637	13.4	78.8	1783	10	AD110403	Adi10403 Human cel
C 565	13.4	78.8	1176	13	ADU67837	Adu67837 DNA encod	C 638	13.4	78.8	1783	11	ADN39842	Adn39842 Cancer/ra
C 566	13.4	78.8	1191	14	ABEB35264	Aeb35264 Human Gef	C 639	13.4	78.8	1783	12	ADJ46927	Adj46927 Human tra
C 567	13.4	78.8	1203	14	ADY20767	Ady20767 Streptomy	C 640	13.4	78.8	1783	13	ADRI18849	Adri18849 Human hep
C 568	13.4	78.8	1203	14	ADY20769	Ady20769 Streptomy	C 641	13.4	78.8	1783	13	ADU06248	Adu06248 Novel bro
C 569	13.4	78.8	1215	11	ABD01346	Abd01346 Klebsiell	C 642	13.4	78.8	1783	14	AEA29003	Aea29003 Phospholi
C 570	13.4	78.8	1215	13	ADQ84944	Adq84944 Human tum	C 643	13.4	78.8	1783	14	AEA29003	Aea29003 Phospholi
C 571	13.4	78.8	1215	13	ADQ84133	Adq84133 Human tum	C 644	13.4	78.8	1798	13	ADX47566	Adx47566 Plant ful
C 572	13.4	78.8	1217	8	ACF34501	Acf34501 Gene enco	C 645	13.4	78.8	1814	2	AAQ88068	Aaq88068 Bacillus
C 573	13.4	78.8	1217	12	ADK51969	Adk51969 Human ato	C 646	13.4	78.8	1814	2	AAV02473	Aav02473 Bacillus
C 574	13.4	78.8	1217	12	ADN05083	Adn05083 Antipsori	C 647	13.4	78.8	1819	8	ABV75877	Abv75877 Streptomy
C 575	13.4	78.8	1217	13	ACN38366	Acn38366 Tumour-as	C 648	13.4	78.8	1823	6	ABK90875	Abk90875 CDNA enco
C 576	13.4	78.8	1217	13	ADP55564	Adp55564 Human PRO	C 649	13.4	78.8	1823	5	ABV29724	Abv29724 Human pro
C 577	13.4	78.8	1217	14	ADY18005	Ady18005 DNA encod	C 650	13.4	78.8	1823	5	ABV29724	Abv29724 Human pro
C 578	13.4	78.8	1217	14	ADY18177	Ady18177 DNA encod	C 651	13.4	78.8	1823	5	ABV23847	Abv23847 Human pro
C 579	13.4	78.8	1221	11	ACH97037	Ach97037 Klebsiell	C 652	13.4	78.8	1823	5	ADU07656	Adu07656 DNA seque
C 580	13.4	78.8	1224	8	ACA45531	Aca45531 Prokaryot	C 653	13.4	78.8	2048	6	ABN87500	Abn87500 SNARE ass
C 581	13.4	78.8	1224	13	ADX64460	Adx64460 Plant ful	C 654	13.4	78.8	2066	10	ADB61902	Adb61902 Human cDN
C 582	13.4	78.8	1288	12	ADQ83653	Adq83653 Human tum	C 655	13.4	78.8	2130	4	AAH15576	Aah15576 Human cDN
C 583	13.4	78.8	1288	12	ADQ86973	Adq86973 Human tum	C 656	13.4	78.8	2130	13	ADR99547	Adr99547 Aoptosis
C 584	13.4	78.8	1288	12	ADQ85809	Adq85809 Human tum	C 657	13.4	78.8	2175	11	ADMO1874	Admo1874 Human cDN
C 585	13.4	78.8	1293	5	ABA17628	Abal7628 Human ner	C 658	13.4	78.8	2175	14	AEC84804	Aec84804 Human cDN
C 586	13.4	78.8	1395	8	ACA31528	Aca31528 Prokaryot	C 659	13.4	78.8	2182	13	ADX63531	Adx63531 Plant ful
C 587	13.4	78.8	1404	14	AEA79693	Aea79693 Cellulomo	C 660	13.4	78.8	2182	12	ADQ63081	Adq63081 Novel hum
C 588	13.4	78.8	1419	8	ABS55927	Abs55927 DNA encod	C 661	13.4	78.8	2234	3	AAA49186	Aaa49186 CDNA enco
C 589	13.4	78.8	1419	8	ABX08453	Abx08453 DNA encod	C 662	13.4	78.8	2252	13	ADQ39082	Adq39082 Human SNP
C 590	13.4	78.8	1419	10	ADG16223	Adg16223 Amylase p	C 663	13.4	78.8	2262	4	AAI61220	Aai61220 Human pol
C 591	13.4	78.8	1419	13	ADU07377	Adu07377 DNA seque	C 664	13.4	78.8	2311	8	ACC46451	Acc46451 Human dit
C 592	13.4	78.8	1425	13	ADSS0219	Adss0219 Bacterial	C 665	13.4	78.8	2326	14	ADY51818	Ady51818 T. rubrum
C 593	13.4	78.8	1449	13	ADSS5733	Adss5733 Bacterial	C 666	13.4	78.8	2363	6	ABL68569	AbL68569 Kidney ca
C 594	13.4	78.8	1455	14	ACL70001	Acl70001 M. xanthu	C 667	13.4	78.8	2363	6	RAD40828	Rad40828 Human hep
C 595	13.4	78.8	1479	8	ABX05056	Abx05056 Human nov	C 668	13.4	78.8	2363	6	RAD40647	Rad40647 Human hep
C 596	13.4	78.8	1482	6	AAAS62259	Aas62259 CDNA sequ	C 669	13.4	78.8	2363	6	ABN97195	Abn97195 Gene #369
C 597	13.4	78.8	1488	14	AEA79692	Aea79692 Cellulomo	C 670	13.4	78.8	2363	12	ABSS59541	Abss59541 Human cDN
C 598	13.4	78.8	1548	2	AAK59634	Aak59634 DNA encod	C 671	13.4	78.8	2363	12	ADL56829	AdL56829 Human NOV
C 599	13.4	78.8	1548	2	AAK59594	Aak59594 Wild type	C 672	13.4	78.8	2363	12	ADO39213	Ado39213 Human cDN
C 600	13.4	78.8	1548	3	AAA48482	Aaa48482 Bacillus	C 673	13.4	78.8	2363	15	AEE49544	Aee49544 Human hep
C 601	13.4	78.8	1548	4	AAA37851	Aaa37851 B. stearo	C 674	13.4	78.8	2364	15	AEEF5103	AeeF5103 Human gen
C 602	13.4	78.8	1548	6	ABL50566	AbL50566 B. stearo	C 675	13.4	78.8	2365	10	ADI22674	Adi22674 Human liv

676	13.4	78.8	2369	13	ADT18820	Adt18820 Plant CDN	c 749	13.4	78.8	9413	10	ADF88596	Adf88596 Hepatitis
677	13.4	78.8	2396	4	AAH14255	Aah14255 Human CDN	750	13.4	78.8	10703	4	ABL02456	AbL02456 Drosophil
c 678	13.4	78.8	2417	4	ABL25400	AbL25400 Drosophil	751	13.4	78.8	10794	15	AEF27799	Aef27799 Human lam
679	13.4	78.8	2434	5	AAS71215	Aas71215 DNA encod	752	13.4	78.8	10800	6	ABQ93881	Abq93881 Human lam
680	13.4	78.8	2582	4	AAK94887	Aak94887 Human ful	753	13.4	78.8	10800	12	AD010046	Ad010046 Novel hum
681	13.4	78.8	2582	12	ADL12057	Adl12057 Full leng	754	13.4	78.8	10800	15	AEF27575	Aef27575 Human lam
682	13.4	78.8	2582	14	ADY19327	Ady19327 DNA encod	755	13.4	78.8	10809	6	ABQ93879	Abq93879 Human lam
c 683	13.4	78.8	2647	4	ABL10486	AbL10486 Drosophil	756	13.4	78.8	10809	12	AD0L0042	Ad0L0042 Novel hum
c 684	13.4	78.8	2652	4	ABL25386	AbL25386 Drosophil	757	13.4	78.8	10809	15	AEF27571	Aef27571 Human lam
685	13.4	78.8	2712	14	ACL63991	AcL63991 M. xanthu	c 758	13.4	78.8	10923	4	ABK42192	AbK42192 Genomic s
c 686	13.4	78.8	2762	10	ADE59805	AdE59805 Human gen	c 759	13.4	78.8	10923	9	ADB60348	AdB60348 Connectiv
c 687	13.4	78.8	2787	6	ABK88156	AbK88156 Alpha-iso	c 760	13.4	78.8	11091	6	AAD27805	Aad27805 Human lam
688	13.4	78.8	2808	14	AEA48187	Aea48187 DNA encod	761	13.4	78.8	11118	6	AAAD27804	AaAd27804 Human lam
c 689	13.4	78.8	2811	10	ADG33162	AdG33162 Human DNA	762	13.4	78.8	11344	13	ACN40816	AcN40816 Tumour-as
c 690	13.4	78.8	2811	12	ADP10639	AdP10639 Reference	763	13.4	78.8	11350	6	ABQ72906	AbQ72906 ECMCAD ge
c 691	13.4	78.8	2821	4	AAD03889	Aad03889 Human tar	764	13.4	78.8	11367	6	AD128066	Ad128066 Human imm
c 692	13.4	78.8	2826	6	ABQ90502	AbQ90502 M. capsul	c 765	13.4	78.8	11425	4	AAK79581	Aak79581 Human imm
c 693	13.4	78.8	2834	12	ADK00580	AdK00580 HOMO prot	766	13.4	78.8	11426	12	ADN04477	AdN04477 Antipsori
694	13.4	78.8	2861	4	AHH18377	Ahh18377 Human CDN	767	13.4	78.8	11426	14	ADY25764	AdY25764 MRAC LAMA
c 695	13.4	78.8	2915	13	ADO87114	AdO87114 Human tum	768	13.4	78.8	11426	14	ADZ48983	AdZ48983 Insulin s
c 696	13.4	78.8	2926	10	ADC30475	AdC30475 Human nov	c 769	13.4	78.8	11438	14	AED73396	Aed73396 Human pla
697	13.4	78.8	2930	10	ADD18909	AdD18909 Human dis	770	13.4	78.8	11640	12	ADM87010	AdM87010 Human pro
c 698	13.4	78.8	2930	10	ADF76392	AdF76392 Novel hum	c 771	13.4	78.8	12292	4	AAS59537	Aas59537 Propionib
699	13.4	78.8	2930	12	ADL91830	AdL91830 Human PRO	c 772	13.4	78.8	12292	4	ACF64466	AcF64466 Propionib
c 700	13.4	78.8	2930	14	AEL82833	Ael82833 Breast ca	c 773	13.4	78.8	18195	14	ACL64716	AcL64716 M. xanthu
c 701	13.4	78.8	2965	4	ABL06205	AbL06205 Drosophil	c 774	13.4	78.8	20001	13	ADM77128	AdM77128 Type II d
c 702	13.4	78.8	3014	10	ADC30474	AdC30474 Human nov	775	13.4	78.8	21779	14	ADM97249	AdM97249 Synthetic
c 703	13.4	78.8	3048	11	ADM02265	AdM02265 Human CDN	776	13.4	78.8	21779	14	AED46581	Aed46581 Synthetic
704	13.4	78.8	3048	14	AEC85195	Aec85195 Human CDN	777	13.4	78.8	21784	10	ABD62371	AbD62371 Human pro
705	13.4	78.8	3070	10	ADC37242	AdC37242 Nuclear f	778	13.4	78.8	21784	10	ABX16008	AbX16008 Human nov
706	13.4	78.8	3108	14	ADY19431	AdY19431 DNA encod	779	13.4	78.8	21784	14	AED64261	Aed64261 Human ser
c 707	13.4	78.8	3275	3	ABL25118	AbL25118 Drosophil	c 780	13.4	78.8	23877	12	ADQ97235	AdQ97235 Human can
c 708	13.4	78.8	3337	1	AAH91089	Aah91089 Fragment	c 781	13.4	78.8	29521	4	AAS32517	Aas32517 Human gen
709	13.4	78.8	3360	14	AEA48185	Aea48185 DNA of th	c 782	13.4	78.8	29543	4	AAS32516	Aas32516 Human gen
710	13.4	78.8	3634	13	ADRO7805	AdR07805 Full leng	c 783	13.4	78.8	29912	4	AAS59507	Aas59507 Propionib
c 711	13.4	78.8	3749	3	ADD24894	AdD24894 DNA encod	c 784	13.4	78.8	29912	8	ACF64436	AcF64436 Propionib
c 712	13.4	78.8	4014	3	AAA13390	AaA13390 Human MT4	785	13.4	78.8	32186	5	ABA16117	AbA16117 Human ner
c 713	13.4	78.8	4014	3	AAA14352	AaA14352 Human MT4	786	13.4	78.8	32192	4	ABK42529	AbK42529 Genomic s
c 714	13.4	78.8	4167	4	RAK84449	Rak84449 Human imm	c 787	13.4	78.8	32192	4	AAAL05532	AaAL05532 Human rep
c 715	13.4	78.8	4388	2	AAT60073	Aat60073 Selective	c 788	13.4	78.8	32192	4	AAAL05257	AaAL05257 Human rep
716	13.4	78.8	4744	4	ABL02457	AbL02457 Drosophil	789	13.4	78.8	32192	4	ABAL05631	AbAL05631 Human tes
717	13.4	78.8	5423	14	ADX06845	AdX06845 Cyclin-de	c 790	13.4	78.8	32192	4	ABL98141	AbL98141 Human ner
718	13.4	78.8	5522	4	ABL06204	AbL06204 Drosophil	c 791	13.4	78.8	32192	5	ABA15328	AbA15328 Human ner
c 719	13.4	78.8	5609	6	AAD23867	Aad23867 PCPC717 D	c 792	13.4	78.8	32192	9	ADB60685	AdB60685 Connectiv
c 720	13.4	78.8	5609	6	AAD23868	Aad23868 PER200 DN	c 793	13.4	78.8	33362	13	ABD32700	AbD32700 Human can
c 721	13.4	78.8	5811	6	ABK88159	AbK88159 DNA encod	794	13.4	78.8	39198	5	AAF58067	Aaf58067 Human pol
c 722	13.4	78.8	5811	8	ADA26477	Ada26477 Alpha-iso	c 795	13.4	78.8	42035	4	AAS59562	Aas59562 Propionib
723	13.4	78.8	5913	4	RAK68376	Rak68376 Human imm	c 796	13.4	78.8	42035	8	ACF64491	AcF64491 Propionib
724	13.4	78.8	5913	4	AAK68731	Aak68731 Human imm	c 797	13.4	78.8	46553	4	AAK81745	Aak81745 Human imm
c 725	13.4	78.8	6153	8	ADA26478	Ada26478 Alpha-iso	798	13.4	78.8	46553	4	AAK67926	Aak67926 Human imm
c 726	13.4	78.8	6196	14	ACL64325	AcL64325 M. xanthu	c 799	13.4	78.8	49442	14	AEB96543	Aeb96543 Human IFN
c 727	13.4	78.8	6270	14	ADU70295	AdU70295 Plasmid p	800	13.4	78.8	49999	2	AAZ23902	Aaz23902 Human LOB
c 728	13.4	78.8	6640	4	ABK42190	AbK42190 Genomic s	c 801	13.4	78.8	49999	6	ABQ09084	AbQ09084 M. capsul
c 729	13.4	78.8	6640	9	ADB60346	AdB60346 Connectiv	c 802	13.4	78.8	57296	4	AAK78847	Aak78847 Human imm
730	13.4	78.8	7477	14	ACL64487	AcL64487 M. xanthu	c 803	13.4	78.8	57296	4	AAK78170	Aak78170 Human imm
731	13.4	78.8	7503	4	AAF80291	Aaf80291 Nucleotid	c 804	13.4	78.8	57296	4	AAK79364	Aak79364 Human imm
732	13.4	78.8	7503	4	AAF80293	Aaf80293 Nucleotid	c 805	13.4	78.8	57296	4	AAK86799	Aak86799 Human imm
733	13.4	78.8	7742	2	AAT84745	Aat84745 Bordetell	806	13.4	78.8	74788	11	ACN45118	AcN45118 Human gen
c 734	13.4	78.8	7989	6	AAD25325	Aad25325 Hepatitis	c 807	13.4	78.8	100990	12	ADJ94407	AdJ94407 Yersinia
735	13.4	78.8	8296	6	ABQ72930	AbQ72930 Human lam	c 808	13.4	78.8	106373	13	ABD32737	AbD32737 Human can
c 736	13.4	78.8	8974	4	ABK42191	AbK42191 Genomic s	c 809	13.4	78.8	122336	13	ABD33303	AbD33303 Human can
c 737	13.4	78.8	8974	9	ADB60347	AdB60347 Connectiv	810	13.4	78.8	149671	6	ABK84797	AbK84797 Moesin CD
738	13.4	78.8	8987	4	AAF80296	Aaf80296 Nucleotid	811	13.4	78.8	149671	9	ADB70361	AdB70361 Human mal
739	13.4	78.8	9390	4	AAF80292	Aaf80292 Nucleotid	812	13.4	78.8	149671	12	ADJ37140	AdJ37140 Human mal
c 740	13.4	78.8	9390	4	AAF80294	Aaf80294 Nucleotid	c 813	13.4	78.8	160921	11	ACN44962	AcN44962 Human gen
c 741	13.4	78.8	9400	13	ADR82191	AdR82191 Hepatitis	814	13.4	78.8	167163	10	ADE82948	AdE82948 Human PVT
c 742	13.4	78.8	9400	13	ADT86648	AdT86648 Hepatitis	c 815	13.4	78.8	167932	10	ADL13501	AdL13501 Osteoarth
c 743	13.4	78.8	9413	2	AAQ80498	Aaq80498 DNA encod	c 816	13.4	78.8	175077	11	ACN44626	AcN44626 Human gen
c 744	13.4	78.8	9413	2	AAQ81559	Aaq81559 Hepatitis	c 817	13.4	78.8	180283	14	AED18452	Aed18452 Fibrotic
c 745	13.4	78.8	9413	6	AAT03960	Aat03960 Partial H	c 818	13.4	78.8	185035	6	ABT10147	AbT10147 Human bre
c 746	13.4	78.8	9413	6	AAD25517	Aad25517 Hepatitis	819	13.4	78.8	185035	8	ACA64951	AcA64951 Human FEN
c 747	13.4	78.8	9413	8	RAAL53723	RaAL53723 Hepatitis	820	13.4	78.8	185035	12	ADQ20284	AdQ20284 Human sof
c 748	13.4	78.8	9413	8	AAD49655	Aad49655 Hepatitis	821	13.4	78.8	191010	12	ADO25291	AdO25291 Human pro

822	13.4	78.8	200418	11	ACN444226	ACN444226 Human gen	c 895	13	76.5	3320	13	ADR06984	ADR06984 Full leng
823	13.4	78.8	215974	12	ADQ97523	Adq97523 Human can	896	13	76.5	4626	12	ADI61636	ADI61636 Human cDN
c 824	13.4	78.8	337344	13	ABd32715	ABd32715 Human can	897	13	76.5	4626	12	AEA43801	AEA43801 Human cDN
c 825	13.4	78.8	349980	5	AAH68532	AAH68532 C Glutami	c 898	13	76.5	4649	10	ADBF69161	ADBF69161 C. neofor
826	13	76.5	20	6	AB197470	AB197470 Capture o	899	13	76.5	5352	10	ADBF81803	ADBF81803 Leukaemia
c 827	13	76.5	24	6	AB191510	AB191510 Capture o	900	13	76.5	5477	4	ABL17951	ABL17951 Drosophil
c 828	13	76.5	24	6	AB191511	AB191511 Capture o	901	13	76.5	5617	4	ABL09445	ABL09445 Drosophil
829	13	76.5	115	4	ABA71199	ABA71199 Human foe	c 902	13	76.5	5729	10	ADF59801	ADF59801 Human con
830	13	76.5	115	4	AAJ51433	AAJ51433 Probe #20	c 903	13	76.5	5729	10	ADSL11395	ADSL11395 Human the
831	13	76.5	115	4	AAK45489	AAK45489 Human bon	c 904	13	76.5	7367	14	AEC04225	AEC04225 Human bre
832	13	76.5	115	4	AAK19498	AAK19498 Human bra	c 905	13	76.5	7397	13	ADS10051	ADS10051 Human the
833	13	76.5	115	4	AB45175	AB45175 Human liv	906	13	76.5	7532	14	AEC04231	AEC04231 Human bre
834	13	76.5	115	6	ABS19757	ABS19757 Human gen	907	13	76.5	7571	14	AEC04227	AEC04227 Human bre
835	13	76.5	183	4	AA848705	AA848705 Pseudomon	c 908	13	76.5	8354	6	ABS52100	ABS52100 Human TEN
836	13	76.5	183	8	ACA15751	ACA15751 Prokaryot	c 909	13	76.5	8354	10	ADF74829	ADF74829 Murine NO
837	13	76.5	201	13	ADS41052	ADS41052 Human aut	c 910	13	76.5	8355	10	ADF74841	ADF74841 Murine NO
838	13	76.5	201	13	ADS40848	ADS40848 Human aut	c 911	13	76.5	8438	6	ABN85378	ABN85378 Human NOV
839	13	76.5	201	13	ADS41054	ADS41054 Human aut	c 912	13	76.5	8624	12	ADQ20349	ADQ20349 Human sof
840	13	76.5	304	6	ABL76352	ABL76352 Corn tass	c 913	13	76.5	8645	6	ABS78652	ABS78652 Human cDN
c 841	13	76.5	383	4	AA837825	AA837825 Novel hum	c 914	13	76.5	8676	4	ABL09444	ABL09444 Drosophil
842	13	76.5	418	6	ABN24111	ABN24111 Human ORF	c 915	13	76.5	8774	12	ADQ24356	ADQ24356 Human sof
843	13	76.5	419	13	ADR62794	ADR62794 Cotton cD	c 916	13	76.5	8907	4	ABL17950	ABL17950 Drosophil
844	13	76.5	426	3	AA828066	AA828066 Human sec	c 917	13	76.5	11221	6	ABS54740	ABS54740 Genomic D
845	13	76.5	465	4	ABA58701	ABA58701 Human foe	918	13	76.5	11726	5	ABA15436	ABA15436 Human ner
846	13	76.5	465	4	AAI38380	AAI38380 Probe #70	919	13	76.5	11726	8	ABZ73812	ABZ73812 Secreted
847	13	76.5	465	4	AAK32555	AAK32555 Human bon	c 920	13	76.5	11726	8	ADA98478	ADA98478 Human sec
848	13	76.5	465	4	AAK06837	AAK06837 Human bra	c 921	13	76.5	16000	12	ADI66880	ADI66880 Human DEX
849	13	76.5	465	4	ABS32266	ABS32266 Human liv	c 922	13	76.5	16579	4	AAI05755	AAI05755 Human rep
850	13	76.5	465	6	ABS07345	ABS07345 Human gen	923	13	76.5	16579	5	ABA20363	ABA20363 Human ner
c 851	13	76.5	474	9	ACH35997	ACH35997 Human end	924	13	76.5	17418	12	ADQ19430	ADQ19430 Human sof
c 852	13	76.5	493	9	ACH25966	ACH25966 Human adu	925	13	76.5	17488	12	ADP13539	ADP13539 Renal cel
c 853	13	76.5	501	4	AA851525	AA851525 Pseudomon	926	13	76.5	19303	6	ABQ99655	ABQ99655 Human mem
c 854	13	76.5	501	8	ACA19459	ACA19459 Prokaryot	927	13	76.5	32733	13	ADS36491	ADS36491 Human aut
c 855	13	76.5	504	11	ABD06150	ABD06150 Pseudomon	928	13	76.5	32733	13	ADS36491	ADS36491 Human aut
c 856	13	76.5	513	11	ABD06203	ABD06203 Pseudomon	929	13	76.5	37286	4	AA859522	AA859522 Proplonib
857	13	76.5	516	5	ADL11454	ADL11454 Human ova	930	13	76.5	54945	8	ACF64451	ACF64451 Proplonib
858	13	76.5	516	5	ADL36612	ADL36612 Human ova	c 931	13	76.5	54945	8	ACF70517	ACF70517 Human sph
859	13	76.5	537	11	ACH96932	ACH96932 Klebsiell	c 932	13	76.5	65563	14	AD212493	AD212493 Murine ca
860	13	76.5	540	6	ABV83565	ABV83565 Human bre	c 933	13	76.5	77522	12	ADS36495	ADS36495 Human aut
c 861	13	76.5	543	10	ABT40345	ABT40345 Toxicity	c 934	13	76.5	104729	12	ADQ18615	ADQ18615 Human sof
c 862	13	76.5	543	13	ADV33260	ADV33260 Rat card	c 935	13	76.5	109646	14	AEI18286	AEI18286 Fibrotic
863	13	76.5	583	11	ADN40640	ADN40640 Rice endo	c 936	13	76.5	110000	4	AAI99682_17	AAI99682_17 Continuation (18 o
864	13	76.5	605	13	ADR62793	ADR62793 Cotton cD	c 937	13	76.5	110000	4	AAI99683_17	AAI99683_17 Continuation (18 o
c 865	13	76.5	637	12	ADJ43465	ADJ43465 Plant cDN	c 938	13	76.5	110000	5	AAI61373_1	AAI61373_1 Continuation (2 of
c 866	13	76.5	654	4	AAH07756	AAH07756 Human cDN	c 939	13	76.5	114771	12	ADQ17641	ADQ17641 Human sof
867	13	76.5	727	6	AA896576	AA896576 Corn prom	c 940	13	76.5	129722	6	ABQ86117	ABQ86117 Human ost
868	13	76.5	746	6	AA896577	AA896577 Corn prom	941	13	76.5	129722	12	ADQ18027	ADQ18027 Human sof
869	13	76.5	799	4	AAH01667	AAH01667 Shewanell	c 942	13	76.5	133893	9	AA854538	AA854538 Human pho
870	13	76.5	809	6	AA896559	AA896559 Corn prom	c 943	13	76.5	163218	15	AEF92730	AEF92730 Human chr
871	13	76.5	828	6	AA896574	AA896574 Corn prom	c 944	13	76.5	342998	14	ABE05055	ABE05055 Cancer-as
872	13	76.5	847	6	AA896575	AA896575 Corn prom	945	12.8	75.3	33	6	ABK47620	ABK47620 Human ser
873	13	76.5	876	13	ADS64193	ADS64193 Bacterial	946	12.8	75.3	36	10	ADB68033	ADB68033 PCR prime
874	13	76.5	888	13	ADS63811	ADS63811 Bacterial	947	12.8	75.3	39	6	ABK81248	ABK81248 Polymun
875	13	76.5	894	13	ADS63443	ADS63443 Bacterial	948	12.8	75.3	39	6	ABK81252	ABK81252 Polymun
876	13	76.5	910	6	AA896560	AA896560 Corn prom	949	12.8	75.3	39	11	ADP83682	ADP83682 HLA-A24 g
c 877	13	76.5	989	4	ABA77099	ABA77099 Prolifera	950	12.8	75.3	39	11	ADL99528	ADL99528 Mouse pol
c 878	13	76.5	1014	10	ADC24088	ADC24088 DNA seque	951	12.8	75.3	39	11	ADL99532	ADL99532 Mouse pol
879	13	76.5	1014	11	ABD06086	ABD06086 Pseudomon	952	12.8	75.3	41	2	AAAX08617	AAAX08617 Primer us
c 880	13	76.5	1014	12	ADH36189	ADH36189 Chemical	953	12.8	75.3	50	4	AAAL31081	AAAL31081 Human SNP
c 881	13	76.5	1014	12	ADG93888	ADG93888 Nitrilase	954	12.8	75.3	50	4	AAAL31082	AAAL31082 Human SNP
c 882	13	76.5	1014	12	ADI62487	ADI62487 DNA encod	955	12.8	75.3	50	4	AAAL31082	AAAL31082 Human SNP
c 883	13	76.5	1014	12	ADI64608	ADI64608 DNA encod	c 956	12.8	75.3	59	2	AAQ53902	AAQ53902 Lambda CII
c 884	13	76.5	1319	11	ACL35553	ACL35553 Rice stre	c 957	12.8	75.3	59	3	AAA12877	AAA12877 Bacteriop
c 885	13	76.5	1467	13	ADS63948	ADS63948 Bacterial	c 958	12.8	75.3	59	9	ACD67228	ACD67228 Trp promo
c 886	13	76.5	1476	13	ADS64316	ADS64316 Bacterial	c 959	12.8	75.3	59	10	ADC34627	ADC34627 Bacteriop
c 887	13	76.5	1503	13	ADS63575	ADS63575 Bacterial	c 960	12.8	75.3	59	10	ADH92051	ADH92051 FGF-nucle
888	13	76.5	1527	4	ABA09025	ABA09025 Human ren	961	12.8	75.3	90	6	ABK36635	ABK36635 HCV DNA e
c 889	13	76.5	1844	10	ADB63459	ADB63459 Human cDN	962	12.8	75.3	188	6	AB872923	AB872923 Human gen
c 890	13	76.5	2000	11	ACL36021	ACL36021 Rice stre	963	12.8	75.3	194	4	AA848750	AA848750 Pseudomon
c 891	13	76.5	2165	12	ADQ63620	ADQ63620 Novel hum	964	12.8	75.3	194	8	ACA15815	ACA15815 Prokaryot
c 892	13	76.5	2244	10	ADB69883	ADB69883 C. neofor	965	12.8	75.3	206	14	ADY53863	ADY53863 Gordonia
c 893	13	76.5	2271	13	ADS60256	ADS60256 Bacterial	966	12.8	75.3	210	3	AA820778	AA820778 Human sec
c 894	13	76.5	2649	10	ADB69522	ADB69522 C. neofor	c 967	12.8	75.3	244	2	AAV54617	AAV54617 LU105 spe

c 968 12.8 75.3 246 2 AAT26025 Human gen
c 969 12.8 75.3 248 7 ADS65052
c 970 12.8 75.3 248 7 ADS65016 Corn seed
c 971 12.8 75.3 249 6 ABL18009 Human ORF
c 972 12.8 75.3 251 2 AAX11441
c 973 12.8 75.3 252 10 ADI16346
c 974 12.8 75.3 254 6 AAD35785
c 975 12.8 75.3 254 14 ADY59580
c 976 12.8 75.3 258 10 ADI36345
c 977 12.8 75.3 259 6 ABL78224 Human ORF
c 978 12.8 75.3 262 6 AAD35783
c 979 12.8 75.3 262 14 ADY59578
c 980 12.8 75.3 263 6 AAD35784
c 981 12.8 75.3 263 14 ADY59579
c 982 12.8 75.3 264 5 ABA09609 Human bon
c 983 12.8 75.3 264 9 ADA47999 Rice gene
c 984 12.8 75.3 269 7 ADS65606
c 985 12.8 75.3 270 4 AAS51110 Salmonell
c 986 12.8 75.3 271 14 ADX38537
c 987 12.8 75.3 288 4 ABL23447 Drosophil
c 988 12.8 75.3 312 10 ADI36344 Full-leng
c 989 12.8 75.3 320 13 ADU76397 DNA probe
c 990 12.8 75.3 325 2 AAH87506 Human sin
c 991 12.8 75.3 330 3 AAC02983 Human sec
c 992 12.8 75.3 330 6 ABL77832 Human kin
c 993 12.8 75.3 346 5 ABA09693 Human bon
c 994 12.8 75.3 366 4 AA191909 Human pol
c 995 12.8 75.3 367 5 ABA11438 Human ner
c 996 12.8 75.3 373 10 ADD27316 Human adi
c 997 12.8 75.3 379 4 AAK55424 Human imm
c 998 12.8 75.3 381 8 ACA24045 Prokaryot
c 999 12.8 75.3 384 3 AAF15257 Trichoder
c1000 12.8 75.3 384 13 ADU59298 Trichoder

ALIGNMENTS

RESULT 1
ADI28528
XX ADI28528 standard; DNA; 17 BP.
AC ADI28528;
DT 22-APR-2004 (first entry)
XX Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA primer.
DE
XX Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; PCR; primer; ss.
XX Homo sapiens.
OS
XX WO2004001060-A2.
FN
XX 31-DEC-2003.
PD
XX 20-JUN-2003; 2003WO-US019255.
PF
XX 20-JUN-2002; 2002US-0390850P.
PR
XX 29-AUG-2002; 2002US-0407006P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Whitney GS, Opitack G, Ramanathan CS, McKinnon M;
PI Bennett KL, Barber LE, Cacace A, Tsuchihasi Z;
XX
XX WPI; 2004-090973/09.
DR
XX New nucleic acid molecule encoding a human G-protein coupled receptor,
PI RAI3, useful for preventing, treating or ameliorating chronic obstructive
PT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
PI

PT of COPD.

XX Disclosure; SEQ ID NO 31; 301pp; English.

XX The present sequence is that of a sense PCR primer based on DNA encoding
XX a peptide ADI28458 from human G-protein coupled receptor retinoic acid
XX induced 3 (RAI-3) ADI28460. The peptide was identified in studies of RAI-
XX 3 phosphorylation in response to cigarette smoke. The primer can be used
XX in a PCR to identify individuals at risk for developing chronic
XX obstructive pulmonary disease (COPD). RAI-3 is tyrosine phosphorylated
XX and/or is associated/complexed with tyrosine phosphorylated proteins only
XX in cells that have been exposed to cigarette smoke. Since RAI-3 is
XX primarily expressed in lung tissue, and since cigarette smoke is a major
XX causative factor of COPD, RAI-3 provides a novel cellular target for
XX identifying modulators, e.g. agonists or antagonists, useful for the
XX treatment and/or prevention of COPD and related disorders such as
XX emphysema and chronic bronchitis. RAI-3 modulators, e.g. agonists and
XX antagonists, especially antisense compounds, can be used to treat COPD
XX and other disorders an diseases associated with regulation of NF-kB
XX and/or its associated or interacting signaling molecules. Single
XX nucleotide polymorphisms (SNPs) detected in the RAI-3 gene are useful for
XX determining COPD association in individuals. RAI-3 nucleic acid molecules
XX and polypeptides are useful for preventing, treating or ameliorating
XX disorders related to aberrant GPCR signaling or cell cycle regulation,
XX pulmonary disorders, inflammatory lung disorders, COPD, the underlying
XX symptoms of COPD, COPD-related disorders or conditions, autoimmune
XX disorders, disorders related to hyperimmune activity, inflammatory
XX conditions, disorders related to aberrant acute phase responses,
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,
XX organ transplant rejection, renal diseases, ischaemia-reperfusion injury,
XX heart disorders, disorders related to aberrant signal transduction,
XX proliferation disorders, cancers, HIV infection, asthma, cystic fibrosis,
XX pulmonary fibrosis, ulcerative colitis, cerebral infarct, myocardial
XX infarct, diabetic nephropathy, allergic rhinitis, Crohn's disease,
XX atherosclerosis, rheumatoid arthritis, inflammatory/autoimmune disorders,
XX glioblastoma, pulmonary small cell undifferentiated carcinoma, carcinoma
XX of the breast, colon, lung, ovary, pancreas, prostate, non-Hodgkin's
XX lymphoma, disorders associated with aberrant cell adhesion, I-CAM
XX function and/or regulation, E-selectin function and/or regulation, or
XX aberrant NF-kB function and/or regulation (all claimed).

XX Sequence 17 BP; 2 A; 7 C; 6 G; 2 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCCAGCTTGGCCGAG 17
|||||
Db 1 GCCCAGCTTGGCCGAG 17

RESULT 2

ADI28527
ID ADI28527 standard; DNA; 40 BP.

XX AC ADI28527;

XX 22-APR-2004 (first entry)

XX Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA.

XX Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; ss.

XX Homo sapiens.

XX WO2004001060-A2.

XX 31-DEC-2003.

XX 20-JUN-2003; 2003WO-US019255.

```
XX 20-JUN-2002; 2002US-0390850P.
PR 29-AUG-2002; 2002US-0407006P.
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX Whitney GS, Opitack G, Garulacan L, Ramanathan CS, McKinnon M;
PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX WPI: 2004-090973/09.
DR P-PSDB; ADI28458.
XX
PT New nucleic acid molecule encoding a human G-protein coupled receptor,
PT RAI3, useful for preventing, treating or ameliorating chronic obstructive
PT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
PT of COPD.
XX
PS Disclosure; SEQ ID NO 30; 301pp; English.
XX
CC The present sequence is that of DNA encoding a peptide corresponding to
CC amino acids 340-353 of a human G-protein coupled receptor, denoted
CC retinoic acid induced 3 (RAI-3) protein ADI28460. The peptide was
CC identified in studies of RAI-3 phosphorylation in response to cigarette
CC smoke. Primers based on this sequence can be used in a PCR to identify
CC individuals at risk for developing chronic obstructive pulmonary disease
CC (COPD). RAI-3 is tyrosine phosphorylated and/or is associated/complexed
CC with tyrosine phosphorylated proteins only in cells that have been
CC exposed to cigarette smoke. Since RAI-3 is primarily expressed in lung
CC tissue, and since cigarette smoke is a major causative factor of COPD,
CC RAI-3 provides a novel cellular target for identifying modulators, e.g.
CC agonists or antagonists, useful for the treatment and/or prevention of
CC COPD and related disorders such as emphysema and chronic bronchitis. RAI-
CC 3 modulators, e.g. agonists and antagonists, especially antisense
CC compounds, can be used to treat COPD and other disorders an diseases
CC associated with regulation of NF-kB and/or its associated or interacting
CC signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
CC the RAI-3 gene are useful for determining COPD association in
CC individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
CC preventing, treating or ameliorating disorders related to aberrant GPCR
CC signaling or cell cycle regulation, pulmonary disorders, inflammatory
CC lung disorders, COPD, the underlying symptoms of COPD, COPD-related
CC disorders or conditions, autoimmune disorders, disorders related to
CC hyperimmune activity, inflammatory conditions, disorders related to
CC aberrant acute phase responses, hypercongenital conditions, birth
CC defects, necrotic lesions, wounds, organ transplant rejection, renal
CC diseases, ischaemia-reperfusion injury, heart disorders, disorders
CC related to aberrant signal transduction, proliferation disorders,
CC cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
CC ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
CC neuropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
CC rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
CC pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
CC regulation (all claimed).
XX
SQ Sequence 40 BP; 11 A; 12 C; 10 G; 7 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 12; Length 40;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGGCCGAG 17
DB 1 GCCCAGCTTGGCCGAG 17
RESULT 3
ADT96418
ID ADT96418 standard; cDNA; 497 BP.
XX
AC ADT96418;
```

```
XX 16-DEC-2004 (first entry)
DT Colon cancer associated human cDNA sequence #1925.
DE
XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX Homo sapiens.
OS
XX US2003087818-A1.
PN
XX 08-MAY-2003.
PD
XX 01-FEB-2002; 2002US-00066543.
PF
XX 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305285P.
PR 16-AUG-2001; 2001US-0313077P.
XX (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX WPI: 2003-040540/03.
XX
DR New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
PT
XX Claim 1; SEQ ID NO 1937; 87pp; English.
PS
XX The invention relates to polynucleotide and polypeptide sequences
XX associated with cancer, particularly colon cancer. Also disclosed are (i)
XX an expression vector comprising the polynucleotide, (ii) a host cell
XX transformed or transfected with the expression vector, (iii) an isolated
XX antibody, or its antigen-binding fragment, which specifically binds to
XX the polypeptide, (iv) a method of detecting or determining the presence
XX of cancer in a patient, (v) a fusion protein comprising at least one of
XX the polypeptides, (vi) an oligonucleotide that hybridises to the
XX polynucleotide sequence under highly stringent conditions, and (vii) a
XX method of stimulating and/or expanding T cells specific for a tumour
XX protein. The polypeptide specifically comprises the amino acid sequence
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
XX frames (ORFs) of C636S. These polypeptides are encoded by the
XX polynucleotide sequences, where both are capable of eliciting a humoral
XX and/or cellular immune response. The polynucleotides, polypeptides, and
XX antibodies are useful for diagnosing, preventing or treating cancer,
XX particularly colon cancer. The polynucleotide and polypeptide sequences
XX are also useful in DNA strand invasion, antisense inhibition, mutational
XX analysis, nucleic acid purification, isolation of transcriptionally
XX active genes, blocking or transcription factor binding, genome cleavage
XX or in situ hybridisation, and as enhancers of transcription or
XX biomarkers. This sequence represents a human colon cancer associated
XX cDNA. Note: The sequence data for this patent was obtained in electronic
XX format directly from the USPTO web site at seqdata.uspto.gov
SQ Sequence 497 BP; 126 A; 130 C; 129 G; 112 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 11; Length 497;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGGCCGAG 17
DB 344 GCCCAGCTTGGCCGAG 360
RESULT 4
```



```
ADX42900
ID ADX42900 standard; cDNA; 497 BP.
XX
AC ADX42900;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1937.
XX
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
XX ss; gene.
XX
OS Homo sapiens.
XX
FN WO200274156-A2.
XX
XX 26-SEP-2002.
XX
FF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
XX
PR 07-FEB-2001; 2001US-0267382P.
XX
PR 11-MAY-2001; 2001US-0290322P.
XX
PR 12-JUL-2001; 2001US-0305265P.
XX
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secríst H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PR New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1937; 244pp; English.
XX
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
XX polypeptides, antibodies are useful for diagnosing, preventing or
XX treating cancer, particularly colon cancer. The nucleic acid and
XX polypeptides are also useful in DNA strand invasion, antisense
XX inhibition, mutational analysis, nucleic acid purification, isolation of
XX transcriptionally active genes, blocking or transcription factor binding,
XX genome cleavage or in situ hybridization, and as enhancers of
XX transcription or biomarkers. The kits are useful for detecting antibody
XX binding. The present sequence represents a human cDNA encoding a colon
XX cancer protein.
XX
SQ Sequence 497 BP; 126 A; 130 C; 129 G; 112 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 11; Length 497;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGCCGAG 17
Db 344 GCCCAGCGTTGCCGAG 360
RESULT 5
ADT96292/c
ID ADT96292 standard; cDNA; 552 BP.
XX
AC ADT96292;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1799.
XX
DE Colon cancer; T cell; tumour protein; C634S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW
```

```
immunostimulant; human; ss.
KW
XX Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
FF 01-FEB-2002; 2002US-00066543.
XX
XX 02-FEB-2001; 2001US-0267400P.
XX
PR 07-FEB-2001; 2001US-0267382P.
XX
PR 11-MAY-2001; 2001US-0290322P.
XX
PR 12-JUL-2001; 2001US-0305265P.
XX
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secríst H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PR New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1811; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
XX associated with cancer, particularly colon cancer. Also disclosed are (i)
XX an expression vector comprising the polynucleotide, (ii) a host cell
XX transformed or transfected with the expression vector, (iii) an isolated
XX antibody, or its antigen-binding fragment, which specifically binds to
XX the polypeptide, (iv) a method of detecting or determining the presence
XX of cancer in a patient, (v) a fusion protein comprising at least one of
XX the polypeptides, (vi) an oligonucleotide that hybridises to the
XX polynucleotide sequence under highly stringent conditions, and (vii) a
XX method of stimulating and/or expanding T cells specific for a tumour
XX protein. The polypeptide specifically comprises the amino acid sequence
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
XX frames (ORFs) of C636S. These polypeptides are encoded by the
XX polynucleotide sequences, where both are capable of eliciting a humoral
XX and/or cellular immune response. The polynucleotides, polypeptides, and
XX antibodies are useful for diagnosing, preventing or treating cancer,
XX particularly colon cancer. The polynucleotide and polypeptide sequences
XX are also useful in DNA strand invasion, antisense inhibition, mutational
XX analysis, nucleic acid purification, isolation of transcriptionally
XX active genes, blocking or transcription factor binding, genome cleavage
XX or in situ hybridisation, and as enhancers of transcription or
XX biomarkers. This sequence represents a human colon cancer associated
XX cDNA. Note: The sequence data for this patent was obtained in electronic
XX format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGCCGAG 17
Db 154 GCCCAGCGTTGCCGAG 138
RESULT 6
ADX42774/c
ID ADX42774 standard; cDNA; 552 BP.
XX
AC ADX42774;
XX
DT 21-APR-2005 (first entry)
XX
```



```
DE Human cDNA encoding colon cancer protein SEQ ID NO 1811.
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
OS
XX Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
XX
PR 07-FEB-2001; 2001US-0267382P.
XX
PR 11-MAY-2001; 2001US-0290322P.
XX
PR 12-JUL-2001; 2001US-0305265P.
XX
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1811; 244pp; English.
XX
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGCCGAG 17
Db 154 GCCACGCTTGCCGAG 138
RESULT 7
AAZ90046
ID AAZ90046 standard; cDNA; 603 BP.
XX
AC AAZ90046;
XX
XX 09-MAY-2000 (first entry)
XX
DE Hydrophobic domain containing protein clone HP10549 coding sequence.
XX
XX Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW anti-inflammatory; infection; bodily characteristic; ss.
XX
OS Homo sapiens.
XX
XX WO200000506-A2.
PN
```

```
XX 06-JAN-2000.
PD
XX 18-JUN-1999; 99WO-JP003242.
XX
XX 26-JUN-1998; 98JP-00180008.
XX
XX (SAGA ) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
PI
XX
XX WPI; 2000-160665/14.
DR
XX P-PSDB; AAY78809.
XX
XX Novel human proteins having hydrophobic domains used for research and
PT diagnostic purposes.
XX
XX Claim 3; Page 88; 117pp; English.
XX
XX This sequence represents the hydrophobic domain containing protein, clone
CC HP10549 coding region. The sequence is isolated from a human stomach
CC cancer cell line. The invention relates to human proteins with
CC hydrophobic domains, the DNA and the cDNA encoding them. The
CC polynucleotides and proteins are predicted to have biological activities
CC which make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals. Suggested activities include
CC nutritional activity (nutritional source or supplement); cytokine and
CC cell proliferation/differentiation activity; immune stimulating (e.g. as
CC vaccines) or suppressing activity (e.g. to treat various immune
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation, Guillain-Barre syndrome, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC organ transplantation); haematopoiesis regulating activity (e.g. in
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or cardiac cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
XX Sequence 603 BP; 133 A; 179 C; 137 G; 154 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 17; DB 3; Length 603;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACGCTTGCCGAG 17
Db 550 GCCACGCTTGCCGAG 566
RESULT 8
ABQ58527/c
ID ABQ58527 standard; cDNA; 620 BP.
XX
XX ABQ58527;
XX
XX 02-AUG-2002 (first entry)
XX
XX Human colon cancer related nucleotide sequence SEQ ID NO:2222.
DE
XX Human colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
KW
```

```
XX OS Homo sapiens.
XX PN W0200229086-A2.
XX XX
XX PD 11-APR-2002.
XX XX
XX PF 02-OCT-2001; 2001WO-US030732.
XX XX
XX PR 02-OCT-2000; 2000US-0237271P.
XX XX
XX PA (FARB ) BAYER CORP.
XX XX
XX PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX PI Thiagalingam A, Lewis ME;
XX DR WPI; 2002-426115/45.
XX XX
XX PF New isolated nucleic acid that is differentially expressed in cancer
XX PR tissues useful for determining the presence of colon cancer in a cell or
XX XX tissue type, and in antisense therapy.
XX PA
XX PS Claim 1; Fig 1; 796pp; English.
XX XX
XX PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX PI Thiagalingam A, Lewis ME;
XX DR WPI; 2002-426115/45.
XX XX
XX PF New isolated nucleic acid that is differentially expressed in cancer
XX PR tissues useful for determining the presence of colon cancer in a cell or
XX XX tissue type, and in antisense therapy.
XX PA
XX PS Claim 1; Fig 1; 796pp; English.
XX XX
XX CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
XX CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX CC used in antisense therapy. An antibody immunoreactive with a polypeptide
XX CC encoded by (I) is useful for detecting cancer in a patient sample, and
XX CC for detecting the presence or absence of a polynucleotide encoded by a
XX CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
XX CC from (I) can be used for determining the presence of a nucleic acid which
XX CC hybridises to (I), and for determining the phenotype of cells in a sample
XX CC of cells from a patient. (I) is useful for determining the presence of
XX CC colon cancer in a cell or tissue type, for determining the presence or
XX CC state of other type of cancer, in antisense therapy, to generate
XX CC macroarrays on a solid surface, to identify a chromosome on which the
XX CC corresponding gene resides, and in tissue profiling, forensics, genetic
XX CC analysis, mapping and diagnostic applications. (I) can be used to raise
XX CC antibodies, and to screen for peptide analogues and antagonists
XX XX
XX SQ Sequence 620 BP; 152 A; 154 C; 161 G; 142 T; 0 U; 11 Other;
XX
XX Query Match 100.0%; Score 17; DB 6; Length 620;
XX Best Local Similarity 100.0%; Pred. No. 64;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Oy 1 GCCCACCCTTGGCCGAG 17
Db 154 GCCCACCCTTGGCCGAG 138
XX
XX RESULT 9
XX ABQ59698/c
XX ID ABQ59698 standard; cDNA; 634 BP.
XX XX
XX AC ABQ59698;
XX XX
XX DT 02-AUG-2002 (first entry)
XX XX
XX DE Human colon cancer related nucleotide sequence SEQ ID NO:3393.
XX XX
XX KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX XX
XX OS Homo sapiens.
XX XX
XX PN W0200229086-A2.
XX XX
XX PD 11-APR-2002.
XX XX
XX PF 02-OCT-2001; 2001WO-US030732.
XX XX
XX PR 02-OCT-2000; 2000US-0237271P.
XX XX
XX PA (FARB ) BAYER CORP.
XX XX
XX PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX PI Thiagalingam A, Lewis ME;
XX DR WPI; 2002-426115/45.
XX XX
XX PF New isolated nucleic acid that is differentially expressed in cancer
XX PR tissues useful for determining the presence of colon cancer in a cell or
XX XX tissue type, and in antisense therapy.
XX PA
XX PS Claim 1; Fig 1; 796pp; English.
XX XX
XX CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
XX CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX CC used in antisense therapy. An antibody immunoreactive with a polypeptide
XX CC encoded by (I) is useful for detecting cancer in a patient sample, and
XX CC for detecting the presence or absence of a polynucleotide encoded by a
XX CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
XX CC from (I) can be used for determining the presence of a nucleic acid which
XX CC hybridises to (I), and for determining the phenotype of cells in a sample
XX CC of cells from a patient. (I) is useful for determining the presence of
XX CC colon cancer in a cell or tissue type, for determining the presence or
XX CC state of other type of cancer, in antisense therapy, to generate
XX CC macroarrays on a solid surface, to identify a chromosome on which the
XX CC corresponding gene resides, and in tissue profiling, forensics, genetic
XX CC analysis, mapping and diagnostic applications. (I) can be used to raise
XX CC antibodies, and to screen for peptide analogues and antagonists
XX XX
XX SQ Sequence 634 BP; 153 A; 154 C; 168 G; 146 T; 0 U; 13 Other;
XX
XX Query Match 100.0%; Score 17; DB 6; Length 634;
XX Best Local Similarity 100.0%; Pred. No. 64;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Oy 1 GCCCACCCTTGGCCGAG 17
Db 158 GCCCACCCTTGGCCGAG 142
XX
XX RESULT 10
XX ABT22328/c
XX ID ABT22328 standard; DNA; 642 BP.
XX XX
XX AC ABT22328;
XX XX
XX DT 16-APR-2003 (first entry)
XX XX
XX DE Breast cancer marker gene SEQ ID No 701.
XX XX
XX KW Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;
XX KW chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;
XX KW surrogate marker gene; pharmacodynamic marker gene; transgenic animal;
XX KW human; ds.
XX XX
XX OS Homo sapiens.
XX XX
XX PN W0200285298-A2.
XX XX
XX PD 31-OCT-2002.
XX XX
XX PF 19-APR-2002; 2002WO-US012612.
XX XX
XX PR 20-APR-2001; 2001US-0285163P.
XX XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX XX
XX PI Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;
XX XX WPI; 2003-093053/08.
XX XX
```

XX Novel isolated polypeptide encoded by breast cancer marker gene, useful
PT for diagnosing, staging, monitoring, prognosing and treating diseases
PT associated with breast cancer.

XX Disclosure; Page 187-188; 725pp; English.

XX The invention relates to an isolated polypeptide encoded by a breast
CC cancer marker gene comprising any of 1417 21-805 nucleotide sequences,
CC given in the specification. The methods of the invention are useful for
CC diagnosing patients having an identified breast cancer or symptoms
CC associated with breast cancer, to diagnose breast cancer or its
CC precursors, and for monitoring the efficacy of treatment of a breast
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also
CC useful for evaluating a patient before, after or during therapy, to
CC evaluate the reduction in a tumour burden. The breast cancer marker gene
CC proteins are useful as immunogens for raising antibodies, by immunising a
CC mammal with a breast cancer marker protein. The marker proteins are
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to
CC identify other proteins which bind to or interact with the marker
CC proteins. The breast cancer marker genes are useful as surrogate marker
CC genes for one or more disorders, disease states or conditions leading to
CC disease states, in particular, breast cancers. The breast cancer marker
CC genes are useful as pharmacodynamic marker genes. An antibody which
CC selectively binds to a protein of a breast cancer marker gene is useful
CC for treating cancers, particularly breast cancers. The host cell of the
CC invention is useful for producing non-human transgenic animals. This
CC polynucleotide sequence represents one of the breast cancer marker genes
CC of the invention

XX Sequence 642 BP; 148 A; 160 C; 180 G; 142 T; 0 U; 12 Other;
SQ

Best Match 100.0%; Score 17; DB 10; Length 642;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCACGCTTGCCGAG 17
|||
Db 184 GCCACGCTTGCCGAG 168

RESULT 11
AD030035
ID AD030035 standard; cDNA; 1074 BP.

XX AC AD030035;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human GPCR RA13 polynucleotide, SEQ ID NO:1137.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrheic;
KW dermatological; antiulcer; antihypertensive; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW gene; ss.

XX Homo sapiens.
XX
XX W02004040000-A2.
XX
XX 13-MAY-2004.
XX

PF 09-SEP-2003; 2003WO-US028226.
XX
PR 09-SEP-2002; 2002US-0409303P.
PR 09-APR-2003; 2003US-0461329P.
XX
XX (PRIM-) PRIMAL INC.
XX
XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX
DR WPI; 2004-390329/36.
XX P-PSDB; ADO29632.
XX
PT Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 1137; 542pp; English.

XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1074 BP; 215 A; 316 C; 259 G; 284 T; 0 U; 0 Other;
SQ

Best Match 100.0%; Score 17; DB 12; Length 1074;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCACGCTTGCCGAG 17
|||
Db 1018 GCCACGCTTGCCGAG 1034

RESULT 12
AAZ61776
ID AAZ61776 standard; cDNA; 1212 BP.
XX
XX AAZ61776;
XX
XX 27-MAR-2000 (first entry)
DT
XX cDNA encoding human skin cell transmembrane protein, SEQ ID NO:249.
DE
XX

KW Skin, dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumour vascularisation; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS Homo sapiens.
XX
XX WO9955865-A1.
XX
XX 04-NOV-1999.
XX
XX 29-APR-1999; 99WO-NZ000051.
XX
XX 29-APR-1998; 98US-00069726.
XX
XX 09-NOV-1998; 98US-00188930.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
XX WPI; 2000-072177/06.
XX
XX Novel polynucleotides useful for the treatment of various conditions
XX including wounds and cancer.
XX
XX Claim 1; Page 155; 235pp; English.
XX
XX The invention relates to novel nucleic acid sequences derived from rat
XX dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
XX and mouse embryonic skin, keratinocyte stem cells and transit amplifying
XX cells. Polypeptides of the invention may be used to treat inflammation,
XX cancer and neurological diseases. The proteins may be used to stimulate
XX the growth and motility of keratinocytes, to inhibit the growth of cancer
XX cells, to modulate angiogenesis and tumour vascularisation, to modulate
XX skin inflammation, to modulate epithelial cell growth and to inhibit
XX binding of HIV-1 to leukocytes. The invention may also be used to treat
XX growth and developmental defects, skin wounds and hair follicle
XX disorders. Sequences AAZ61606-261832 represent cDNA sequences derived
XX from several mouse, rat or human skin cell types. Sequences AAZ61606-
XX Z61649, AAZ61725-261765, AAZ61802-261811 and AAZ61826 encode proteins
XX with an N-terminal signal sequence, indicating that the proteins are
XX secreted. Sequences AAZ61650-261668, AAZ61766-261780, AAZ61812-261817 and
XX AAZ61827-261829 encode proteins with one or more putative transmembrane
XX domains
XX
SQ Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCGTTGGCCGAG 17
Db 990 GCCCAGCGTTGGCCGAG 1006
RESULT 13
AAC99709
ID AAC99709 standard; cDNA; 1212 BP.
XX
XX AAC99709;
XX
XX 08-MAR-2001 (first entry)
XX
XX Skin cell cDNA, SEQ ID NO: 249.
XX
XX Human; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;
XX neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease; ss.
XX

OS Homo sapiens.
XX
XX WO200069884-A2.
XX
XX 23-NOV-2000.
XX
XX 15-MAY-2000; 2000WO-NZ000075.
XX
XX 14-MAY-1999; 99US-00312283.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;
XX Murison JG;
XX
XX WPI; 2001-007495/01.
XX
XX P-PSDB; AAB55958.
XX
XX New isolated polynucleotide used in the identification of genetic
XX disorders and encoding polypeptides used for treating inflammatory
XX disease, cancer and neurological diseases.
XX
XX Claim 1; Page 210-211; 352pp; English.
XX
XX The present polynucleotide encodes a polypeptide which is expressed in
XX mammalian skin cells. The polypeptide is useful for stimulating
XX keratinocyte growth and motility, inhibiting the growth of cancer cells,
XX modulating angiogenesis, inhibiting angiogenesis and vascularisation of
XX tumours, modulating skin inflammation, stimulating the growth of
XX epithelial cells, inhibiting the binding of human immunodeficiency virus
XX (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
XX neurological diseases. The polynucleotide can be used as a marker, in the
XX identification of genetic disorders, and for the design of
XX oligonucleotides for examining expression patterns
XX
SQ Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCGTTGGCCGAG 17
Db 990 GCCCAGCGTTGGCCGAG 1006
RESULT 14
ABL34861
ID ABL34861 standard; cDNA; 1212 BP.
XX
XX ABL34861;
XX
XX 04-APR-2002 (first entry)
XX
XX Human cDNA isolated from skin cells SEQ ID NO: 249.
XX
XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
XX developmental defect; inflammatory disease; dermatological; vulnery;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;
KW ss.
XX
XX Homo sapiens.
XX
XX WO200190357-A1.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-NZ000099.
XX
XX 24-MAY-2000; 2000US-0206650P.
XX
XX 25-JUL-2000; 2000US-0221232P.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;
PI Kumble KD;
XX WPI; 2002-122020/16.
XX New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.
XX Claim 1; Page 175; 466pp; English.
XX The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC cDNA of the invention
XX
SQ Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGCCGAG 17
Db 990 GCCCAGCGTTGCCGAG 1006
RESULT 15
ACA04775
ID ACA04775 standard; cDNA; 1228 BP.
AC ACA04775;
XX
XX 28-MAY-2003 (first entry)
XX cDNA encoding human membrane associated protein fragment #223.
XX Human; ss; gene; microarray; membrane-associated protein; neuropathology;
KW immunopathology; pancreatic disease; cancer; diabetes; hyperlipidaemia;
KW pancreatic cholera; Alzheimer's disease; Huntington's disease; sarcoma;
KW fibrocystic disease; leukaemia; adenocarcinoma; AIDS; allergy; anaemia;
KW asthma; gout; dementia.
XX Homo sapiens.
XX US6492505-B1.
PN
XX 10-DEC-2002.
PD
XX 31-JAN-2000; 2000US-00495050.
XX
XX 01-FEB-1999; 99US-0118318P.
XX (INCYTE) INCYTE GENOMICS INC.
XX
XX Reddy R, Guegler KJ, Au-Young J;
XX WPI; 2003-327324/31.
XX
XX Combination for research/diagnostic applications and for monitoring
PT treatment of e.g., cancer, comprises polynucleotides comprising a
PT fragment of gene encoding membrane-associated proteins, receptors or ion
PT channels.
XX
XX Claim 1; Col 215-216; 147pp; English.

CC The invention relates to a combination comprising several polynucleotide
CC sequences comprising a fragment of gene encoding membrane-associated
CC proteins, receptors or ion channels. The combination is useful as a
CC probe, for research and diagnostic applications, for monitoring the
CC expression of several expressed polynucleotides, in the diagnosis and
CC monitoring of treatment of pancreatic disease, cancer, immunopathology or
CC neuropathology, for investigating an individual's predisposition to the
CC above disease, in genetic or gene expression analysis of polynucleotide
CC sequences, to investigate cellular responses to infection or drug
CC treatment, as hybridisable array elements in a microarray, to purify a
CC subpopulation of mRNAs, cDNAs or genomic fragments in a sample, in
CC diagnostics, prognostics and treatment regimens, in drug discovery and
CC development, in toxicological and carcinogenicity studies, and in
CC forensics or pharmacogenomics, to monitor the progression of disease, to
CC monitor the efficacy of treatment, to diagnose the conditions of the
CC pancreas e.g. diabetes, pancreatic cholera, hyperlipidaemia or
CC fibrocystic disease, to diagnose a cancer e.g. leukaemia, adenocarcinoma
CC or sarcoma, to diagnose immunopathologies e.g. AIDS, allergies, anaemia,
CC asthma or gout, to diagnose neuropathologies e.g. Alzheimer's disease,
CC dementia or Huntington's disease, to rapidly screen large numbers of
CC candidate drug molecules and as query sequences against GenBank.
CC SwissProt, BLOCKS and PRINTS databases. The combination is employed to
CC fine tune the treatment regimen and thus the expression patterns
CC associated with undesirable side effects are avoided. The present
CC sequence represents a cDNA encoding a fragment of gene encoding human
CC membrane-associated proteins, receptors or ion channels
XX
SQ Sequence 1228 BP; 272 A; 341 C; 283 G; 332 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 8; Length 1228;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGCCGAG 17
Db 541 GCCCAGCGTTGCCGAG 557
RESULT 16
AEA00112
ID AEA00112 standard; cDNA; 1460 BP.
XX
XX AEA00112;
XX
XX 28-JUL-2005 (first entry)
DT
XX Human TAT143 cDNA SEQ ID NO:64.
DE
XX ss; gene; tumor-associated antigen; cytostatic; breast tumor;
KW endocrine disease; gynecology and obstetrics; neoplasm; colon tumor;
KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
KW ovary tumor; skin tumor; liver tumor.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 128..1201
FT /*tag= a
FT /product= "TAT143"
XX
XX US2005106644-A1.
PN
XX 19-MAY-2005.
PD
XX 08-SEP-2004; 2004US-00936626.
XX
XX 20-JUN-2001; 2001US-0299500P.
XX 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323288P.
PR 19-JUN-2002; 2002US-0017488.
PR 26-MAR-2004; 2004US-0557116P.
PR 04-AUG-2004; 2004US-0598899P.

XX	(GETH) GENENTECH INC.
PA	
XX	Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
PI	Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI	Slikowski M;
XX	WPI; 2005-384304/39.
DR	P-PSDB; AEA00190.
XX	
XX	Novel isolated antibody capable of binding to tumor-associated antigenic
PT	target polypeptide, useful for treating cell proliferative disorder e.g.
PT	cancer.
XX	
XX	Claim 1; SEQ ID NO 64; 337pp; English.
XX	
CC	The invention relates to a novel isolated antibody binding to a
CC	polypeptide having at least 80% sequence identity to a polypeptide having
CC	any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
CC	target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
CC	specification, a polypeptide having any one of (AEA00127-AEA00202),
CC	lacking its associated signal peptide, or an extracellular domain of a
CC	polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
CC	encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
CC	An antibody of the invention has cytostatic activity. The antibody is
CC	useful for inhibiting growth of a cell expressing TAT188, which involves
CC	contacting the cell with the antibody. The cell is a cancer cell chosen
CC	from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
CC	liver cell. The cancer cell is a mammalian cell, preferably a human cell.
CC	The antibody is also useful for detecting the level of TAT188
CC	polypeptide expressed in a test cell relative to a control cell, and for
CC	detecting the level of TAT188 polypeptide or a polypeptide having at
CC	least 80 % sequence identity to the TAT188 polypeptide sequence in a test
CC	cell relative to a control cell. The antibody is useful for treating a
CC	preventing a cell proliferative disorder associated with increased
CC	expression or activity of a polypeptide having at least 80 % identity to
CC	a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
CC	The method of the invention is useful for inhibiting the growth of a
CC	cancer cell. The present sequence encodes a polypeptide of the invention.
XX	
SQ	Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;
	Query Match 100.0%; Score 17; DB 14; Length 1460;
	Best Local Similarity 100.0%; Pred. No. 65;
	Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GCCACGCTTGCGCAG 17
Db	1145 GCCACGCTTGCGCAG 1161
	RESULT 17
ID	AEA00632
ID	AEA00632 standard; cDNA; 1460 BP.
XX	
AC	AEA00632;
XX	
DT	28-JUL-2005 (first entry)
XX	
DE	Human TAT143 cDNA sequence SeqID64.
XX	
KW	antibody identification; tumor-associated antigen; cytostatic;
KW	RNA interference; gene therapy; cell death; cancer; breast tumor;
KW	colon tumor; renal tumor; lung tumor; liver tumor; ovary tumor;
KW	skin tumor; liver tumor; gene; ss; TAT143.
XX	
OS	Homo sapiens.
XX	
PN	US2005107595-A1.
XX	
PD	19-MAY-2005.
XX	
PF	10-SEP-2004; 2004US-00938061.

XX	20-JUN-2001; 2001US-0299500P.
PR	23-JUN-2001; 2001US-0301880P.
PR	18-SEP-2001; 2001US-0323268P.
PR	19-JUN-2002; 2002US-00177488.
PR	26-MAR-2004; 2004US-0557116P.
PR	04-AUG-2004; 2004US-0598899P.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
PI	Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI	Sakanaka C, Chunharapai A, Reed CJ;
XX	
XX	WPI; 2005-371577/38.
XX	P-PSDB; AEA00710.
XX	
PT	Novel isolated antibody e.g. anti-El6 or anti-TAT112 antibody that binds
PT	to tumor-associated antigenic target polypeptide, useful for diagnosing
PT	or treating cancer.
XX	
PS	Claim 1; SEQ ID NO 64; 96pp; English.
XX	
CC	This invention relates to a novel isolated antibody, for example anti-El6
CC	or anti-TAT112 antibody, that binds to a tumor-associated antigenic
CC	target polypeptide (TAT) and that lacks an associated signal peptide
CC	sequence. The invention may be useful for the development of compounds
CC	with a cytostatic activity acting as antagonists of the TAT188
CC	polypeptide or RNA interference whilst the disclosed sequences may be
CC	useful for gene therapy. The invention is useful for inducing the death
CC	of a cell (such as a cancer cell chosen from breast, colon, rectum,
CC	endometrium, kidney, lung, ovary, skin and liver) to which it binds,
CC	inhibiting proliferation or promoting cell death of a cell expressing
CC	TAT188. In addition, the invention may be useful for detecting the level
CC	of TAT188 polypeptide in a test cell relative to a control cell, or
CC	treatng or preventing a cell proliferative disorder associated with
CC	increased expression of TAT188. The novel antibody of the invention is
CC	useful for inhibiting the growth of a cancer cell and may be useful for
CC	diagnosing or treating cancer. The present sequence is that of the human
CC	TAT143 cDNA which encodes a protein against which an antibody of the
CC	invention may be targeted.
XX	
SQ	Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;
	Query Match 100.0%; Score 17; DB 14; Length 1460;
	Best Local Similarity 100.0%; Pred. No. 65;
	Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GCCACGCTTGCGCAG 17
Db	1145 GCCACGCTTGCGCAG 1161
	RESULT 18
ID	AAF58615
ID	AAF58615 standard; cDNA; 1619 BP.
XX	
AC	AAF58615;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Human RECAP polynucleotide, SEQ ID NO: 43.
XX	
KW	Human; RECAP; receptors and associated proteins; cerebroprotective;
KW	nootropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW	antiidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW	antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW	cytostatic; antibacterial; virucide; fungicide; protozoacide;
KW	antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.
OS	Homo sapiens.
XX	
PN	WO200107612-A2.

```

XX PD 01-FEB-2001.
XX PF
XX PR 21-JUL-2000; 2000WO-US020035.
XX XX
XX PR 21-JUL-1999; 99US-0145232P.
XX PR 07-OCT-1999; 99US-0158578P.
XX PR 12-NOV-1999; 99US-0165192P.
XX XX
XX PA (INCY-) INCYTE GENOMICS INC.
XX XX
XX PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
XX PI Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;
XX XX
XX DR WPI; 2001-168554/17.
XX DR P-PSDB; AAB68891.
XX XX
XX PT Novel receptors and associated proteins for diagnosis and treatment of
XX PT neurological disorders, immunological disorders including autoimmune/
XX PT inflammatory disorders and cell proliferative disorders such as cancer.
XX PS Example 5; Page 126-127; 128pp; English.
XX XX
XX CC The present sequence encodes a human RECAP (receptors and associated
XX CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
XX CC in the diagnosis, treatment and prevention of neurological disorders such
XX CC as stroke, Alzheimer's disease, Pick's disease, Huntington's disease,
XX CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral
XX CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD
XX CC (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker
XX CC syndrome), immunological disorders, including autoimmune/inflammatory
XX CC disorders such as AIDS, DiGeorge's syndrome, severe combined
XX CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Crohn's disease,
XX CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
XX CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,
XX CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
XX CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
XX CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
XX CC cirrhosis, hepatitis and cancer
XX XX
XX SQ Sequence 1619 BP; 331 A; 463 C; 400 G; 425 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 4; Length 1619;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAG 17
DB 1137 GCCACGCTTGCCGAG 1153

RESULT 19
AAZ90056
ID AAZ90056 standard; cDNA; 1718 BP.
XX AC
XX AC AAZ90056;
XX XX
XX DT 09-MAY-2000 (first entry)
XX XX
XX DE Hydrophobic domain containing protein clone HP10549 nucleotide sequence.
XX KW Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
XX KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;
XX KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
XX KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
XX KW anti-inflammatory; infection; bodily characteristic; ss.
XX OS
XX OS Homo sapiens.
XX PN W0200000506-A2.
XX XX
XX PD 06-JAN-2000.
XX XX

```

```

PF 18-JUN-1999; 99WO-JP003242.
XX XX
XX PR 26-JUN-1998; 98JP-00180008.
XX XX
XX PA (SAGA ) SAGAMI CHEM RES CENT.
XX PA (PROT-) PROTEGENE INC.
XX PI Kato S, Kimura T;
XX XX
XX DR WPI; 2000-160665/14.
XX DR P-PSDB; AAY78809.
XX XX
XX PT Novel human proteins having hydrophobic domains used for research and
XX PT diagnostic purposes.
XX PS Claim 4; Page 111-113; 117pp; English.
XX XX
XX CC This sequence represents the hydrophobic domain containing protein, clone
XX CC HP10549 nucleotide sequence. The sequence is isolated from a human
XX CC stomach cancer cell line. The invention relates to human proteins with
XX CC hydrophobic domains, the DNA and the cDNA encoding them. The
XX CC polynucleotides and proteins are predicted to have biological activities
XX CC which make them suitable for treating, preventing or ameliorating medical
XX CC conditions in humans and animals. Suggested activities include
XX CC nutritional activity (nutritional source or supplement); cytokine and
XX CC cell proliferation/differentiation activity; immune stimulating (e.g. as
XX CC vaccines) or suppressing activity (e.g. to treat various immune
XX CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic
XX CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
XX CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
XX CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
XX CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
XX CC organ transplantation); haematopoiesis regulating activity (e.g. in
XX CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
XX CC activity (e.g. wound healing and tissue repair, ulcers, burns,
XX CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
XX CC activity; haemostatic and thrombolytic activity (e.g. treating
XX CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
XX CC tumour inhibition activity. The polynucleotides are also stated to be
XX CC useful for gene therapy. Other activities include inhibiting infections
XX CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
XX CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
XX CC skin; effecting biorhythms or cardiac cycles; enhancing fertility;
XX CC treatment of depression; treatment of pain; hormonal or endocrine
XX CC activity. The polynucleotides may also be used for recombinant expression
XX CC of the protein
XX XX
XX SQ Sequence 1718 BP; 381 A; 495 C; 377 G; 465 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 3; Length 1718;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCCGAG 17
DB 561 GCCACGCTTGCCGAG 577

RESULT 20
ADF70573
ID ADF70573 standard; DNA; 1788 BP.
XX AC
XX AC ADF70573;
XX XX
XX DT 12-FEB-2004 (first entry)
XX XX
XX DE Orphan receptor ligand-related human protein gene SegID196.
XX KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
XX KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
XX KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.
XX OS
XX OS Homo sapiens.

```



```
XX PN WO2003071272-A1.
XX PD 28-AUG-2003.
XX PX 21-FEB-2003; 2003WO-JP001901.
XX PR 22-FEB-2002; 2002JP-00045728.
XX PR 23-JUL-2002; 2002JP-00213949.
XX PR 11-OCT-2002; 2002JP-00298237.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
XX DR WPI; 2003-697654/66.
XX DR P-PSDB; ADF70471.
XX PT Transformation of cells with a fusion protein of an orphan receptor
XX PT protein with a fluorescent protein useful for identification of ligands
XX PT to the orphan receptor.
XX PS Example 4; SEQ ID NO 196; 594pp; Japanese.
XX CC This invention relates to a novel method of identifying ligands to an
XX CC orphan receptor protein which comprises transforming cells with DNA
XX CC encoding a fusion protein of the orphan receptor with a fluorescent
XX CC protein, so that the fusion protein is expressed in the cells for cell
XX CC membranes isolated from them) and contacting the cells with the potential
XX CC ligand to be tested. A suitable fluorescent protein for incorporation in
XX CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
XX CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
XX CC identification of ligands binding to an orphan receptor protein.
XX SQ Sequence 1788 BP; 444 A; 473 C; 412 G; 459 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 10; Length 1788;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGCCGAG 17
Db 1018 GCCCAGCGTTGCCGAG 1034
RESULT 21
ABT10173
ID ABT10173 standard; cDNA; 2302 BP.
AC ABT10173;
XX DT 04-DEC-2002 (first entry)
XX DE Human breast cancer associated coding sequence SEQ ID NO: 307.
XX KW Human; breast specific gene; breast cancer; differential expression;
XX KW cytostatic; gene therapy; gene; ss.
XX OS Homo sapiens.
XX PN WO200259271-A2.
XX PD 01-AUG-2002.
XX PF 25-JAN-2002; 2002WO-US002176.
XX PR 25-JAN-2001; 2001US-0263757P.
XX PR 25-APR-2001; 2001US-0286090P.
XX PR 23-MAY-2001; 2001US-0292517P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Orr MS, Nation M, Diggins JC, Zeng W;
```

```
XX WPI; 2002-674803/72.
XX DR Diagnosing breast cancer in a patient comprises detecting the level of
XX PT gene expression in cell or tissue samples, where a differential gene
XX PT expression is indicative of breast cancer.
XX PS Claim 1; SEQ ID NO 307; 260pp + Sequence Listing; English.
XX CC The present invention relates to methods of diagnosing breast cancer in a
XX CC patient, which comprise detecting the level of expression in a tissue
XX CC sample of two or more genes selected from those shown in ABT09867-
XX CC AB11112, where a differential expression of the genes indicates breast
XX CC cancer. The methods are useful in diagnosing, treating, detecting the
XX CC progression, and in monitoring treatment of breast cancer in patients.
XX CC The methods are also useful as a screening tool for agents that modulate
XX CC the onset or progression of breast cancer. The breast cancer genes may be
XX CC used as diagnostic markers for the prediction or identification of the
XX CC malignant state of breast tissue, for confirming the type and progression
XX CC of cancer, and for drug screening and assays. The present sequence is a
XX CC coding sequence of the invention. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub.published_pct_sequences
XX SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 6; Length 2302;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGCCGAG 17
Db 1117 GCCCAGCGTTGCCGAG 1133
RESULT 22
ACC59386
ID ACC59386 standard; cDNA; 2302 BP.
XX AC ACC59386;
XX DT 26-AUG-2003 (first entry)
XX DE Human GPCR-like retinoic acid-induced gene 1.
XX KW Human; retinoic acid-induced gene 1; RAIG1; feeding; fasting; GPCR;
XX KW receptor; G-protein coupled receptor; anorectic; antidiabetic;
XX KW antidepressant; immunomodulator; transgenic; gene therapy; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 100..1173
XX FT /*tag= a
XX FT /product= "GPCR-like RAIG1"
XX PN WO2003016553-A2.
XX PD 27-FEB-2003.
XX PF 20-AUG-2002; 2002WO-US026510.
XX PR 20-AUG-2001; 2001US-0313940P.
XX PA (GETH ) GENENTECH INC.
XX PA (CURA-) CURAGEN CORP.
XX PI Lewin DA, Stewart TA;
XX DR WPI; 2003-278580/27.
XX DR P-PSDB; ABR42649.
```

PT New G-protein coupled receptor-like retinoic acid induced gene 1 (GPCR-
PT like RAIG1) polypeptide and gene, useful for diagnosing or treating
PT metabolic disorders, e.g. obesity, anorexia, cachexia or diabetes.

XX Claim 10; Page 16-17; 150pp; English.

CC The present sequence is that of human G-protein coupled receptor-like
CC retinoic acid induced gene 1 (GPCR-like RAIG1). This is the human
CC homologue of murine GPCR-like RAIG1. The murine gene was shown to be
CC differentially regulated during fasting-feeding cycles in mice, with
CC moderate induction early in fasting, down-regulation with extended
CC fasting and 4-fold up-regulation with feeding in recovery from fasting.
CC The differentially expressed gene, its mRNA, and the encoded protein, can
CC each be manipulated to detect and treat metabolic disorders associated
CC with up- or down-regulation of GPCR-like RAIG1 activity, such as obesity,
CC anorexia, cachexia or diabetes

XX Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 8; Length 2302;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAG 17

Db 1117 GCCCAGCTTGCCGAG 1133

RESULT 23

ADD93240

ID ADD93240 standard; cDNA; 2302 BP.

XX AC

ADD93240;

XX DT

29-JAN-2004 (first entry)

XX DE

RAIG1 coding sequence.

XX KW

ss; gene; retinoic acid-inducible gene 1; RAIG1; orphan;

XX KW

G-protein coupled receptor; GPCR; chromosome 12; GPCR5B; carcinoma;

XX KW

vaccine; breast cancer; pancreatic cancer; lung cancer; liver cancer;

XX KW

ovarian cancer; colon cancer; osteosarcoma.

XX OS

Homo sapiens.

XX FH

Key Location/Qualifiers

XX CDS

100..1173

FT

/*tag= a

FT

/product= "RAIG1"

XX PN

WO2003087832-A2.

XX PD

23-OCT-2003.

XX PF

10-APR-2003; 2003WO-GB001587.

XX PR

11-APR-2002; 2002GB-00008331.

XX PR

17-SEP-2002; 2002GB-00021538.

XX PA

(OXFO-) OXFORD GLYSCSCIENCES UK LTD.

XX PI

Terrett JA;

XX DR

WPI; 2003-845382/78.

XX DR

P-PSDB; ADD93239.

XX DR

GENBANK; AF095448.

XX PT

Screening, diagnosing and/or treating carcinoma, including breast,

PT pancreatic, lung, liver, ovarian and colon cancer by detecting the change

PT in expression or activity of an RAIG1 polypeptide or encoding nucleic

PT acid molecule.

XX PS

Claim 1; Fig 2; 43pp; English.

XX

CC This sequence encodes retinoic acid-inducible gene 1 (RAIG1) polypeptide.

CC RAIG1 is an orphan G-protein coupled receptor (GPCR) located on

CC chromosome 12. RAIG1 shows a restricted expression pattern compared to a

CC related receptor, GPCR5B, which is widely expressed in peripheral and

CC central tissues. The RAIG1 cDNA and polypeptide sequence may be used in

CC the method of the invention for screening for and/or diagnosis of

CC carcinoma in a subject, and/or monitoring the effectiveness of carcinoma

CC therapy. The method comprises detecting and/or quantifying in a

CC biological sample obtained from the subject an RAIG1 polypeptide and a

CC nucleic acid molecule. The RAIG1 polypeptide and nucleic acid molecule

CC are useful in the manufacture of a medicament for the treatment of

CC carcinoma, where the composition is a vaccine. An agent which interacts

CC with or causes change in the expression or activity of an RAIG1

CC polypeptide or nucleic acid molecule, is also useful in the manufacture

CC of a medicament for the treatment of carcinoma that is breast cancer,

CC pancreatic cancer, lung cancer, liver cancer, ovarian cancer, colon

CC cancer and/or osteosarcoma. They can also be used in the diagnosis and

CC screening of such carcinomas.

XX Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 10; Length 2302;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAG 17

Db 1117 GCCCAGCTTGCCGAG 1133

RESULT 24

ADL24773

ID ADL24773 standard; DNA; 2302 BP.

XX AC

ADL24773;

XX DT

20-MAY-2004 (first entry)

XX DE

Intestinal epithelium/peyer's patch M cell-associated DNA sequence #171.

XX KW

intestinal epithelium cell development; peyer's patch M cell development;

XX KW

inflammatory bowel disease; glutenenteropathy; infectious disease;

XX KW

autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;

XX KW

Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;

XX KW

immune system disorder; hypersensitivity; anaphylaxis;

XX OS

blood group incompatibility; ds; human.

XX OS

Homo sapiens.

XX PN

WO200280852-A2.

XX PD

17-OCT-2002.

XX PF

04-APR-2002; 2002WO-US010873.

XX PR

04-APR-2001; 2001US-0281416P.

XX PA

(DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX PI

Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;

XX WPI; 2003-075470/07.

XX Novel isolated or purified polypeptide encoded by genes associated with

XX intestinal epithelium or M cell development, differentiation or function,

XX useful for treating autoimmune diseases and infectious diseases.

XX Claim 1; SEQ ID NO 283; 152pp; English.

XX The invention comprises DNA sequences which are associated with

XX intestinal epithelium and peyer's patch M cells. The DNA sequences of the

XX invention are useful for assessing, modifying, modulating or regulating

CC intestinal epithelium or M cell development. The DNA sequences of the
CC invention are also useful in the treatment of: inflammatory bowel
CC disease, glutenenteropathy, infectious diseases, autoimmune diseases
CC (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's
CC disease, multiple sclerosis, allergy, asthma and diabetic mellitus),
CC diseases or disorders of the immune system, hypersensitivity,
CC anaphylaxis, and blood group incompatibility. The present nucleic acid
CC represents an intestinal epithelium/peyer's patch M cell-associated DNA
CC sequence of the invention. NOTE: The present sequence is not shown in the
CC specification, but has been retrieved from the WIPO website.

XX Sequence 2302 BP; 494 A; 566 C; 533 G; 509 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 1117 GCCCAGCGTTGCCGAG 1133

RESULT 25

ABT31923
ID ABT31923 standard; DNA; 2316 BP.

XX AC ABT31923;

XX DT 01-MAY-2003 (first entry)

XX DE Human breast cancer / ovarian cancer related coding sequence #30.

XX KW Human; gene; ds; cytostatic; breast cancer; ovarian cancer.

XX OS Homo sapiens.

XX PN WO2003000012-A2.

XX PD 03-JAN-2003.

XX PF 21-JUN-2002; 2002WO-US019773.

XX PR 21-JUN-2001; 2001US-0300159P.

XX PR 27-JUN-2001; 2001US-0301351P.

XX PA (MTLL-) MILLENNIUM PHARM INC.

XX PI Veiby OP;

XX WPI; 2003-267848/26.

XX P-PSDB; ABJ37054.

XX PT Determining the presence of breast cancer in an individual, involves
XX using specific polynucleotide markers.

XX PS Disclosure; Page 163; 233pp; English.

XX CC The invention comprises a method for assessing whether a patient is
XX afflicted with breast cancer or ovarian cancer. The method involves the
XX use of specific DNA markers. The method of the invention is useful in the
XX detection and treatment of ovarian and breast cancer. DNA sequences
XX CC ABT31894 - ABT31949 encode human breast/ovarian cancer-related proteins

XX SQ Sequence 2316 BP; 502 A; 670 C; 535 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 10; Length 2316;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 1123 GCCCAGCGTTGCCGAG 1139

RESULT 26

AAH14688
ID AAH14688 standard; cDNA; 2446 BP.

XX AC AAH14688;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:12388.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EPI074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.

XX PS Claim 8; SEQ ID NO 12388; 2537pp + Sequence Listing; English.

XX CC The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention

XX SQ Sequence 2446 BP; 507 A; 723 C; 572 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 4; Length 2446;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 1271 GCCCAGCGTTGCCGAG 1287


```
CC pathologies
XX
SQ Sequence 2456 BP; 522 A; 720 C; 571 G; 643 T; 0 U; 0 Other;
  Query Match 100.0%; Score 17; DB 10; Length 2456;
  Best Local Similarity 100.0%; Pred. No. 66;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1271 GCCCAGCGTTGGCCGAG 1287

RESULT 29
ADN39301
ID ADN39301 standard; cDNA; 2456 BP.
XX
AC ADN39301;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:619.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2002US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
XX
DR P-PSDB; ADN39302.
XX
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 619; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC
```

```
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
  Query Match 100.0%; Score 17; DB 11; Length 2456;
  Best Local Similarity 100.0%; Pred. No. 66;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1271 GCCCAGCGTTGGCCGAG 1287

RESULT 30
ADI28525
ID ADI28525 standard; cDNA; 2456 BP.
XX
AC ADI28525;
XX
DT 22-APR-2004 (first entry)
XX
DE Human GPCR retinoic acid induced 3 (RAI-3) cDNA.
XX
KW Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiac; antiallergic;
KW broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
KW chromosome 12p13-p12.3; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(112,r)
FT /*tag= a
FT /label= RAI-3-s1
FT /note= "located in 5' untranslated region"
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 254..1327
FT /*tag= b
FT /product= "Human RAI3"
FT variation replace(364,y)
FT /*tag= c
FT /label= RAI-3-s2
FT /note= "exon 1, silent (Ala/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(511,y)
FT /*tag= d
FT /label= RAI-3-s3
FT /note= "exon 2, silent (Ile/Ile)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(523,y)
FT /*tag= e
FT /label= RAI-3-s4
FT /note= "exon 2, silent (Asp/Asp)"
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(605,r)
FT /*tag= f
FT /label= RAI-3-s6
FT /note= "exon 2, missense (Ser/Gly)"
FT
```

```
FT /standard_name= "Single nucleotide polymorphism"
FT replace(797,r)
FT /*tag= g
FT /label= RAI-3-s5
FT /note= "exon 2, missense (Thr/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1111,r)
FT /*tag= h
FT /label= RAI-3-s8
FT /note= "silent (Pro/Pro)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1173,r)
FT /*tag= i
FT /label= RAI-3-s9
FT /note= "missense (Gln/Arg)"
FT /standard_name= "Single nucleotide polymorphism"
XX WO2004001060-A2.
PN
XX
XX 31-DEC-2003.
PD
XX
XX 20-JUN-2003; 2003WO-US019255.
PF
XX
XX 20-JUN-2002; 2002US-0390850P.
PR
XX 29-AUG-2002; 2002US-0407006P.
PR
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA
XX
XX Whitney GS, Opitck G, Garulacan L, Ramanathan CS, McKinnon M;
PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX
XX WPI; 2004-090973/09.
DR
XX P-PSDB; ADI28526.
DR
XX
XX New nucleic acid molecule encoding a human G-protein coupled
XX receptor, RAI3, useful for preventing, treating or ameliorating chronic obstructive
XX pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
XX of COPD.
XX
XX Claim 21; SEQ ID NO 18; 301pp; English.
XX
XX The present sequence is that of cDNA encoding a human G-protein coupled
XX receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
XX Proteomic methods were used to isolate cigarette smoke-inducible
XX tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
XX identified as being tyrosine phosphorylated and/or as being
XX associated/complexed with tyrosine phosphorylated proteins only in those
XX cells that had been exposed to cigarette smoke. Since RAI-3 is primarily
XX expressed in lung tissue, and since cigarette smoke is a major causative
XX factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
XX novel cellular target for identifying modulators, e.g. agonists or
XX antagonists, useful for the treatment and/or prevention of COPD and
XX related disorders such as emphysema and chronic bronchitis. RAI-3
XX modulators, e.g. agonists and antagonists, especially antisense
XX compounds, can be used to treat COPD and other disorders and diseases
XX associated with regulation of NF-kB and/or its associated or interacting
XX signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
XX the RAI-3 gene are useful for determining COPD association in
XX individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
XX preventing, treating or ameliorating disorders related to aberrant GPCR
XX signaling or cell cycle regulation, pulmonary disorders, inflammatory
XX lung disorders, COPD, the underlying symptoms of COPD, COPD-related
XX disorders or conditions, autoimmune disorders, disorders related to
XX hyperimmune activity, inflammatory conditions, disorders related to
XX aberrant acute phase responses, hypercongenital conditions, birth
XX defects, necrotic lesions, wounds, organ transplant rejection, renal
XX diseases, ischaemia-reperfusion injury, heart disorders, disorders
XX related to aberrant signal transduction, proliferation disorders,
XX cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
XX ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
XX nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
XX rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
XX pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
```

```
CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
CC regulation (all claimed).
XX
SQ Sequence 2456 BP; 516 A; 719 C; 570 G; 643 T; 0 U; 8 Other;
Query Match 100.0%; Score 17; DB 12; Length 2456;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGCGCGAG 17
Db 1271 GCCCAGCTTGCGCGAG 1287
RESULT 31
ADI28459
ID ADI28459 standard; cDNA; 2456 BP.
XX
XX ADI28459;
XX
XX 22-APR-2004 (first entry)
XX
XX Human GPCR retinoic acid induced 3 (RAI-3) cDNA.
XX
XX Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
XX antiinflammatory; immunosuppressive; cytosatic; cardiac; antiallergic;
XX broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
XX chromosome 12p13-p12.3; ss.
XX
XX Homo sapiens.
XX
XX Location/Qualifiers
FH replace(112,a)
FT /*tag= a
FT /label= RAI-3-s1
FT /note= "located in 5' untranslated region"
FT /standard_name= "Single nucleotide polymorphism"
FT 254..1327
FT CDS
FT /*tag= b
FT /product= "Human RAI3"
FT replace(364,t)
FT /*tag= c
FT /label= RAI-3-s2
FT /note= "exon 1, silent (Ala/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(511,t)
FT /*tag= d
FT /label= RAI-3-s3
FT /note= "exon 2, silent (Ile/Ile)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(523,t)
FT /*tag= e
FT /label= RAI-3-s4
FT /note= "exon 2, silent (Asp/Asp)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(605,g)
FT /*tag= f
FT /label= RAI-3-s6
FT /note= "exon 2, missense (Ser/Gly)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(797,g)
FT /*tag= g
FT /label= RAI-3-s5
FT /note= "exon 2, missense (Thr/Ala)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1111,g)
FT /*tag= h
FT /label= RAI-3-s8
FT /note= "silent (Pro/Pro)"
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1173,g)
```

```
FT FT /*tag= i
FT FT /label= RAI-3-s9
FT FT /note= "missense (Gln/Arg)"
FT FT /standard_name= "Single nucleotide polymorphism"
XX PN WO2004001060-A2.
XX XX 31-DEC-2003.
XX XX 20-JUN-2003; 2003WO-US019255.
XX XX 20-JUN-2002; 2002US-0390850P.
XX XX 29-AUG-2002; 2002US-0407006P.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX XX Whitney GS, Opitck G, Garulacan L, Ramanathan CS, McKinnon M;
XX PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX XX WPI; 2004-090973/09.
XX DR P-PSDB; ADI28460.
XX DR GENBANK; NM_003979.
XX XX New nucleic acid molecule encoding a human G-protein coupled receptor,
XX FT RAI3, useful for preventing, treating or ameliorating chronic obstructive
XX FT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
XX FT of COPD.
XX PS Claim 1; SEQ ID NO 2; 301pp; English.
XX CC The present sequence is that of cDNA encoding a human G-protein coupled
XX CC receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
XX CC Proteomic methods were used to isolate cigarette smoke-inducible
XX CC tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
XX CC identified as being tyrosine phosphorylated and/or as being
XX CC associated/complexed with tyrosine phosphorylated proteins only in those
XX CC cells that had been exposed to cigarette smoke. Since RAI-3 is primarily
XX CC expressed in lung tissue, and since cigarette smoke is a major causative
XX CC factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
XX CC novel cellular target for identifying modulators, e.g. agonists or
XX CC antagonists, useful for the treatment and/or prevention of COPD and
XX CC related disorders such as emphysema and chronic bronchitis. RAI-3
XX CC modulators, e.g. agonists and antagonists, especially antisense
XX CC compounds, can be used to treat COPD and other disorders and diseases
XX CC associated with regulation of NF-kB and/or its associated or interacting
XX CC signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
XX CC the RAI-3 gene are useful for determining COPD association in
XX CC individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
XX CC preventing, treating or ameliorating disorders related to aberrant GPCR
XX CC signaling or cell cycle regulation, pulmonary disorders, inflammatory
XX CC lung disorders, COPD, the underlying symptoms of COPD, COPD-related
XX CC disorders or conditions, autoimmune disorders, disorders related to
XX CC hyperimmune activity, inflammatory conditions, disorders related to
XX CC aberrant acute phase responses, hypercongenital conditions, birth
XX CC defects, necrotic lesions, wounds, organ transplant rejection, renal
XX CC diseases, ischaemia-reperfusion injury, heart disorders, disorders
XX CC related to aberrant signal transduction, proliferation disorders,
XX CC cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
XX CC ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
XX CC nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
XX CC rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
XX CC pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
XX CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
XX CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
XX CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
XX CC regulation (all claimed).
XX SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 12; Length 2456;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 13; Length 2456;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCGAG 17
|||||
DB 1271 GCCCAGCGTTGCCGCGAG 1287

RESULT 33
ACN39388
ID ACN39388 standard; cDNA; 2456 BP.
XX
AC ACN39388;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA226771, SEQ ID NO:3498.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
WPI: 2004-347921/32.
XX
P-PSDB; ABM81354.
XX
New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 3498; 7273pp; English.
XX
The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acids and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention
```

```
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 17; DB 13; Length 2456;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCGAG 17
|||||
DB 1271 GCCCAGCGTTGCCGCGAG 1287

RESULT 34
ADR43992
ID ADR43992 standard; DNA; 2456 BP.
XX
AC ADR43992;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human breast tumour associated gene clone-02 SEQ ID NO:2.
XX
KW cancer; tumour; breast cancer; lung cancer; colon cancer; kidney cancer;
KW antibody; cytostatic; gene therapy; human; breast tumour tissue;
KW breast tissue; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2004074506-A2.
XX
PD 02-SEP-2004.
XX
PF 30-JAN-2004; 2004WO-US003030.
XX
PR 13-FEB-2003; 2003US-0447900P.
XX
PA (MERG-) MERGEN LTD.
XX
PI Hu Q, Peng A, Liu B, Love JR, Hao X, Ren M, Sheng Z;
XX
WPI; 2004-635589/61.
XX
Diagnosing and treating a cancer, e.g. breast, lung, colon, or kidney
XX cancer, comprises detecting and blocking the over expression of a gene of
XX a protein found in breast, lung, colon, or kidney tissue.
XX
Claim 2; SEQ ID NO 2; 220pp; English.
XX
The present invention describes a method for diagnosing and treating a
XX cancer, e.g. breast, lung, colon, or kidney cancer. The method comprises
XX detecting and blocking the over expression of a gene of a protein found
XX in breast, lung, colon, or kidney tissue. Also described is an antibody
XX or binding portion of an antibody that specifically binds a protein found
XX in breast tissue, lung tissue, colon tissue, or kidney tissue. The
XX antibody has cytostatic activity, and can be used in gene therapy. The
XX methods, antibodies, polynucleotides and polypeptides from the present
XX invention are useful for detecting, diagnosing, preventing and treating
XX cancer, e.g. breast, lung, colon, or kidney cancer. The present sequence
XX represents a nucleotide sequence given in the present invention, which is
XX over expressed in human breast tumour tissue. N.B. All 395 sequences
XX referenced in this patent are detailed in the US provisional application
XX SN 60/447,900, filed 02/13/2003. In this application only one sequence,
XX representing the longest sequence of each of the 65 clones is listed in
XX Appendix A (and given as SEQ ID NO:1 to 65 in the Sequence Listing).
```

```
XX
SQ Sequence 2456 BP; 517 A; 710 C; 567 G; 643 T; 0 U; 19 Other;
Query Match 100.0%; Score 17; DB 13; Length 2456;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCGAG 17
|||||
```

```
Db      1271 GCCCAGCGTTGCCGAG 1287

RESULT 35
ADU06126
ID      ADU06126 standard; DNA; 2456 BP.
XX
AC      ADU06126;
XX
DT      27-JAN-2005 (first entry)
XX
DE      Novel bronchial cancer-associated human gene SeqID350.
XX
KW      bronchial cancer; cytostatic; tumour-associated protein;
KW      cancer detection; metastasis; tumour; gene; ds; human.
XX
OS      Homo sapiens.
XX
PN      DE10316701-A1.
XX
PD      04-NOV-2004.
XX
PF      09-APR-2003; 2003DE-01016701.
XX
PR      09-APR-2003; 2003DE-01016701.
XX
PA      (HINZ/) HINZMANN B.
PA      (HERM/) HERMANN K.
PA      (CAST/) HEIDEN CASTANOS-VELEZ E.
XX
PI      Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
PI      Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pillarsky C;
XX
WPI; 2004-786403/78.
DR      P-PSDB; ADU06613.
XX
PT      New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT      cancer and in screening for therapeutic and diagnostic agents.
XX
PS      Claim 1; SEQ ID NO 350; 1381pp; German.
XX
CC      This invention relates to a novel isolated nucleic acid associated with
CC      bronchial cancer comprising 489 defined sequences given in the
CC      specification. The invention may be useful for the production of
CC      compounds with a cytostatic activity through the inhibition of expression
CC      or activity of tumour-associated proteins. The novel DNA sequences and
CC      the proteins/peptides encoded by them are used for detecting bronchial
CC      cancer or determining the risk of developing it and to screen for
CC      specific binding partners of the DNA or protein sequences, where the
CC      binding partners are potentially useful as agents for treating or
CC      diagnosing bronchial cancer. The DNA or protein sequences can also be
CC      used for prognosis, detection of metastases and for secondary treatment
CC      (of tumours that have been stabilised or are no longer detectable).
CC      Detecting abnormal expression of the DNA sequences provides early
CC      diagnosis of bronchial cancers. The present sequence is that of a novel
CC      bronchial cancer-associated human gene sequence of the invention.
XX
SQ      Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match      100.0%; Score 17; DB 13; Length 2456;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 GCCCAGCGTTGCCGAG 17
      |||||
Db      1271 GCCCAGCGTTGCCGAG 1287

RESULT 36
AEA00088
ID      AEA00088 standard; cDNA; 2456 BP.
XX
AC      AEA00088;
```

Query Match	Best Local Similarity	100.0%;	Score 17;	DB 14;	Length 2456;
Matches	17;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	Indels	0;
Qy	1	GCCACGCTTGGCCGAG 17			
Db	1271	GCCACGCTTGGCCGAG 1287			
RESULT 37					
AEA00608					
ID	AEA00608	standard; cDNA; 2456 BP.			
AC	AEA00608;				
XX					
DT	28-JUL-2005	(first entry)			
XX					
DE	Human TAT115 cDNA sequence SeqID40.				
XX					
KW	antibody identification; tumor-associated antigen; cytostatic;				
KW	RNA interference; gene therapy; cell death; cancer; breast tumor;				
KW	colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;				
KW	skin tumor; liver tumor; gene; ss; TAT115.				
XX					
OS	Homo sapiens.				
XX					
PN	US2005107595-A1.				
XX					
PD	19-MAY-2005.				
XX					
PF	10-SEP-2004; 2004US-00938061.				
XX					
PR	20-JUN-2001; 2001US-0299500P.				
PR	29-JUN-2001; 2001US-0301880P.				
PR	18-SEP-2001; 2001US-0323268P.				
PR	19-JUN-2002; 2002US-00177488.				
PR	26-MAR-2004; 2004US-0557116P.				
PR	04-AUG-2004; 2004US-0598899P.				
XX					
PA	(GETH) GENENTECH INC.				
XX					
PI	Cairns B, Chen R, Prantz G, Hillan KJ, Koeppen H, Phillips HS;				
PI	Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;				
PI	Sakanaka C, Chuncharapai A, Reed CJ;				
XX					
DR	WPI; 2005-371577/38.				
XX	P-PSDB; AEA00686.				
XX					
PT	Novel isolated antibody e.g. anti-El6 or anti-TAT112 antibody that binds				
PT	to tumor-associated antigenic target polypeptide, useful for diagnosing				
PT	or treating cancer.				
XX					
PS	Claim 1; SEQ ID NO 40; 96pp; English.				
XX					
CC	This invention relates to a novel isolated antibody, for example anti-El6				
CC	or anti-TAT112 antibody, that binds to a tumor-associated antigenic				
CC	target polypeptide (TAT) and that lacks an associated signal peptide				
CC	sequence. The invention may be useful for the development of compounds				
CC	with a cytostatic activity acting as antagonists of the TAT188				
CC	polypeptide or RNA interference whilst the disclosed sequences may be				
CC	useful for gene therapy. The invention is useful for inducing the death				
CC	of a cell (such as a cancer cell chosen from breast, colon, rectum,				
CC	endometrium, kidney, lung, ovary, skin and liver) to which it binds,				
CC	inhibiting proliferation or promoting cell death of a cell expressing				
CC	TAT188. In addition, the invention may be useful for detecting the level				
CC	of TAT188 polypeptide in a test cell relative to a control cell, or				
CC	treating or preventing a cell proliferative disorder associated with				
CC	increased expression of TAT188. The novel antibody of the invention is				
CC	useful for inhibiting the growth of a cancer cell and may be useful for				
CC	diagnosing or treating cancer. The present sequence is that of the human				
CC	TAT115 cDNA which encodes a protein against which an antibody of the				
CC	invention may be targeted.				
XX					
SQ	Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;				

PT Substance that suppresses expression of retinoic acid-inducible G-protein
 PT coupled receptor-3 RAI3 gene or nucleic acid, useful for suppressing cell
 PT growth, and as pharmaceutical or therapeutic agent of breast cancer.

XX
 XX
 PS Claim 1; SEQ ID NO 2; 22pp; Japanese.

XX
 CC The invention relates to a novel substance which suppresses the
 CC expression of a retinoic acid-inducible G-protein coupled receptor (RAI)3
 CC gene. The RAI3 gene suppressing agent is useful for suppressing cell
 CC growth, and as a pharmaceutical or therapeutic agent of breast cancer,
 CC and is also useful in identifying a cell growth inhibitory substance. The
 CC RAI3 gene suppressing agent enables the identification of a cell growth
 CC inhibitory substance. This polynucleotide sequence represents the
 CC retinoic acid-inducible G-protein coupled receptor 3 gene of the
 CC invention.

XX
 SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 14; Length 2456;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

Db 1271 GCCCAGCTTGGCCGAG 1287

RESULT 39
 ABQ54954

ID ABQ54954 standard; cDNA; 2593 BP.

XX AC ABQ54954;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HSKJC61 cDNA, SEQ ID NO:834.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; gene; ss.

XX OS Homo sapiens.

XX WO200200677-A1.

XX PN 03-JAN-2002.

XX PD 07-JUN-2001; 2001WO-US018569.

XX PF 07-JUN-2000; 2000US-0209467P.

XX PR (HUMA-) HUMAN GENOME SCI INC.

XX PA Birse CE, Rosen CA;

XX PI WPI; 2002-147878/19.

XX DR P-PSDB; ABP41877.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.

XX PS Claim 1; SEQ ID NO 834; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-

CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2593 BP; 591 A; 730 C; 605 G; 667 T; 0 U; 0 Other;

Query Match 100.0%; Score 17; DB 6; Length 2593;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

Db 1383 GCCCAGCTTGGCCGAG 1399

RESULT 40

AAH72766

ID AAH72766 standard; cDNA; 3371 BP.

XX AC AAH72766;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 4040.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000WO-US033312.

XX PR 08-DEC-1999; 99US-0169681P.

XX PR 21-DEC-1999; 99US-0171350P.

XX PR 14-MAR-2000; 2000US-0189315P.

XX PR 12-MAY-2000; 2000US-0203791P.

XX PR 09-JUN-2000; 2000US-0210600P.

XX PR 21-JUL-2000; 2000US-0220114P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 840; 1051pp; English.

CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 3371 BP; 741 A; 902 C; 779 G; 946 T; 0 U; 3 Other;

Query Match 100.0%; Score 17; DB 4; Length 3371;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
DB 1184 GCCCAGCTTGGCCGAG 1200

RESULT 41
ACN89274
ID ACN89274 standard; DNA; 4239 BP.

XX AC ACN89274;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 10424.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.

XX Disclosure; SEQ ID NO 10424; 36pp; English.

CC The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX Sequence 4239 BP; 859 A; 1161 C; 1075 G; 1141 T; 0 U; 3 Other;

Query Match 100.0%; Score 17; DB 11; Length 4239;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
DB 1679 GCCCAGCTTGGCCGAG 1695

RESULT 42
ADX97494
ID ADX97494 standard; DNA; 6730 BP.

XX AC ADX97494;

XX 21-APR-2005 (first entry)

XX Pancreatic cancer associated human protein encoding DNA, SEQ ID 42.

XX pancreas tumor; cytostatic; gene; ds.

XX Homo sapiens.

XX EP1471075-A2.

XX 27-OCT-2004.

XX 31-MAR-2004; 2004EP-00090124.

XX 31-MAR-2003; 2003DE-01015834.

XX (HINZ/) HINZMANN B.

XX (ROSE/) ROSENTHAL A.

XX (PILA/) PILARSKY C.

XX (DAHL/) DAHL E.

XX (SPEC/) SPECHT T.

XX (LICH/) LICHTNER R.

XX Rosenthal A, Pilarsky C, Dahl E, Specht T, Bruemendorf T;

XX Lichtner R, Staub E, Roepcke S, Li X;

XX WPI; 2004-768082/76.

XX P-PSDB; ADX97565.

XX New nucleic acid differentially expressed in pancreatic tumor tissue, for
XX use as diagnostic agents and in screening for therapeutic agents.

XX Claim 1; SEQ ID NO 42; 28pp; German.

CC The invention relates to a novel human nucleic acid sequence of the
CC pancreas and its encoded protein. The invention further comprises:
CC proteins and peptides, preferably isolated, that contain a sequence
CC encoded by the novel nucleic acid; and methods for diagnosis and
CC treatment of pancreatic cancer, using a substance that inhibits or binds
CC to the protein or its DNA, including: an antisense oligonucleotide, short
CC interfering RNA or ribozyme directed against the pancreatic protein, an
CC organic molecule, particularly having a molecular weight below 5000,
CC especially 300, that binds to the pancreatic DNA, an aptamer or
CC (monoclonal) antibody, preferably human or humanized, that binds to the
CC pancreatic DNA, or an anti-idiotypic antibody raised against the
CC monoclonal antibody, any of which may be derivatized with a reporter
CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
CC human pancreatic proteins and their encoding DNA have cytostatic
CC activity. The novel sequences are useful for inhibiting transcription
CC and/or expression of genes and proteins associated with pancreatic
CC cancer. This polynucleotide sequence represents the DNA encoding one of
CC the novel human pancreatic proteins of the invention. Note: This sequence
CC is not shown in the specification, it has been electronically downloaded
CC from a DVD-ROM provided with this specification by the European Patent
CC Office.

XX Sequence 6730 BP; 1469 A; 1763 C; 1555 G; 1943 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 17; DB 13; Length 6730;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCAGCGTTGGCCGAG 17
      |||||
Db      1271 GCCCAGCGTTGGCCGAG 1287

RESULT 43
ABD12717
ID ABD12717 standard; DNA; 1527 BP.
XX AC ABD12717;
XX XX
DT 29-JUL-2004 (first entry)
DE Pseudomonas aeruginosa polynucleotide #11321.
XX KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX KW antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX XX
PD 22-APR-2003.
XX XX
PF 18-FEB-1999; 99US-00252991.
XX XX
PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR P-PSDB; ABO79146.
XX XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 11321; 455pp; English.
XX CC The invention relates to pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biochip technology. Sequences ABD01397.
XX CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 1527 BP; 266 A; 557 C; 475 G; 229 T; 0 U; 0 Other;

Query Match      90.6%; Score 15.4; DB 11; Length 1527;
Best Local Similarity 94.1%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCCCAGCGTTGGCCGAG 17
      |||||
Db      25 GCCCAGCGTTGGCCGAG 41

RESULT 45
AEA48186
ID AEA48186 standard; DNA; 2976 BP.
XX AC AEA48186;
XX XX
DT 11-AUG-2005 (first entry)
DE DNA of signal peptide and precursor of serine protease, SEQ ID NO: 2.
XX XX
```

```
RESULT 44
ABQ70591/c
ID ABQ70591 standard; DNA; 2306 BP.
XX AC ABQ70591;
XX XX
DT 29-AUG-2003 (revised)
DT 29-AUG-2002 (first entry)
XX XX
DE Listeria monocytogenes 4b contig DNA sequence #533.
XX KW Antibacterial; Listeria; food contamination; mutational analysis;
XX KW infection; ds.
XX OS Listeria monocytogenes ATCC 19115.
XX PN WO200228891-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-FR003061.
XX PR 04-OCT-2000; 2000FR-00012697.
XX XX
XX (INSP ) INST PASTEUR.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PI Kunst F, Glaser P;
XX WPI; 2002-332479/37.
XX XX
PT New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX PS Claim 14; SEQ ID NO 3404; 180pp; French.
XX CC The present invention relates to nucleic acid sequences (ABQ67188-
XX CC ABQ711212) from Listeria sp. The sequences are useful as probes and
XX CC primers for identification and/or detection of Listeria (e.g. as
XX CC contaminants in foods, or mutational analysis) and for analysis of gene
XX CC expression. Proteins encoded by the nucleic acid sequences can be used to
XX CC screen for compounds that modulate gene expression, replication and
XX CC pathogenicity of Listeria (potential therapeutic agents), also for
XX CC treating infections by Listeria, and are useful as immunogens in anti-
XX CC Listeria vaccines. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
XX CC on 29-AUG-2003 to standardise OS field)
XX SQ Sequence 2306 BP; 805 A; 411 C; 489 G; 601 T; 0 U; 0 Other;

Query Match      90.6%; Score 15.4; DB 6; Length 2306;
Best Local Similarity 94.1%; Pred. No. 4.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCCCAGCGTTGGCCGAG 17
      |||||
Db      834 GCCCAGCGTTGGCTGAG 818

RESULT 45
AEA48186
ID AEA48186 standard; DNA; 2976 BP.
XX AC AEA48186;
XX XX
DT 11-AUG-2005 (first entry)
DE DNA of signal peptide and precursor of serine protease, SEQ ID NO: 2.
XX XX
```

KW serine protease; hydrolysis; animal feed; detergent; textile; leather;
 KW ds; gene.
 XX Cellulomonas.
 XX WO2005052161-A2.
 PN 09-JUN-2005.
 PD 19-NOV-2004; 2004WO-US039006.
 XX 19-NOV-2003; 2003US-0523609P.
 XX (GENV) GENENCOR INT INC.
 XX Jones BE, Kolkman M, Leeflang C, Poulouse AJ, Shaw A;
 PI Van Der Kleij WA, Van Marrewijk L;
 XX WPI; 2005-425198/43.
 DR New serine protease (isolated from a member of the Micrococaceae) useful
 PT in e.g. cleaning composition and animal feed composition.
 XX Claim 53; SEQ ID NO 2; 333pp; English.
 XX The present invention relates to new isolated variant serine protease as
 CC given in SEQ ID NO:8 (I), obtained from a member of the Micrococaceae.
 CC Also claimed are compositions comprising an isolated serine protease
 CC having immunological cross-reactivity with the serine proteases obtained
 CC from a member of the Micrococaceae and particularly the protease
 CC obtained from Cellulomonas 69B4. The variant has improved stability as
 CC compared to wild-type Cellulomonas 69B4 protease. The protease is
 CC obtained from Cellulomonas, Oerskovia, Cellulosimicrobium,
 CC Xylanthacterium or from Micromonospora (preferably Cellulomonas 69B4). The
 CC variant serine protease comprises at least one substitution corresponding
 CC to the amino acid positions in SEQ ID NO. 8, and where the variant
 CC protease has better performance in at least one property of keratin
 CC hydrolysis, thermostability, casein activity, LAS stability or cleaning,
 CC as compared to wild-type Cellulomonas 69B4 protease. The invention deals
 CC with serine proteases, genetic material encoding the proteases,
 CC proteolytic proteins obtained from Micrococaceae spp, variant proteins
 CC developed from them, vectors comprising the DNA encoding the protease,
 CC host cells transformed with the vector DNA and enzymes produced by the
 CC host cells. (I) is useful in cleaning compositions and animal feed
 CC compositions, and is useful in laundry and dish detergents. It is useful
 CC in textile and leather processing compositions. The isolated
 CC polynucleotide of (I) provides the capability of isolating further
 CC polynucleotides, which encode proteins having serine protease activity.
 CC The enzyme compositions have comparable or improved wash performance, as
 CC compared to presently used subtilisin proteases. The present sequence is
 CC DNA encoding the signal peptide that is operatively linked to the
 CC precursor protease of serine protease, SEQ ID NO:2.
 XX
 SQ Sequence 2976 BP; 410 A; 1078 C; 1078 G; 410 T; 0 U; 0 Other;
 Query Match 90.6%; Score 15.4; DB 14; Length 2976;
 Best Local Similarity 94.1%; Pred. No. 4.4e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCCACGCTTGCCGAG 17
 DB 239 GCCACGCTTGCCGAG 255
 RESULT 46
 AAX81946
 ID AAX81946 standard; DNA; 4183 BP.
 XX
 AC AAX81946;
 XX
 DT 09-SEP-1999 (first entry)
 XX
 DE DNA encoding Human nucleotide pyrophosphohydrolase-2.

XX Human; nucleotide pyrophosphohydrolase-2; NTPPH-2; arthropathy;
 KW Behcet's syndrome; degenerative joint disease; lupus erythematosus;
 KW rheumatoid arthritis; psoriasis; immunological disease;
 KW acquired immune deficiency syndrome; Addison's disease;
 KW ulcerative colitis; osteoporosis; microbial infection; cancer; ss.
 XX Homo sapiens.
 OS
 XX WO9932610-A1.
 PN 01-JUL-1999.
 PD 02-DEC-1998; 98WO-US025558.
 XX 22-DEC-1997; 97US-00996083.
 PR (INCY-) INCYTE PHARM INC.
 XX Magna H, Schaffer P, Lawton M, Yocum SA, Mitchell PG;
 PI Hutchinson N, Murry LE;
 XX WPI; 1999-418919/35.
 DR P-PSDB; AAY23754.
 XX Human nucleotide pyrophosphohydrolase-2.
 PT Claim 8; Fig 1A-K; 89pp; English.
 XX The present sequence encodes human nucleotide pyrophosphohydrolase-2
 CC (NTPPH-2). Antagonists of the protein are used to treat arthropathy (e.g.
 CC Behcet's syndrome, degenerative joint disease, lupus erythematosus,
 CC rheumatoid arthritis, psoriasis and many others); immunological disease
 CC (e.g. acquired immune deficiency syndrome, Addison's disease, ulcerative
 CC colitis, osteoporosis, microbial infections and many others), also a wide
 CC range of cancers
 XX
 SQ Sequence 4183 BP; 674 A; 1427 C; 1351 G; 728 T; 0 U; 3 Other;
 Query Match 90.6%; Score 15.4; DB 2; Length 4183;
 Best Local Similarity 94.1%; Pred. No. 4.4e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCCACGCTTGCCGAG 17
 DB 3254 GCCACGCTTGCCGAG 3270
 RESULT 47
 ADQ24610
 ID ADQ24610 standard; DNA; 4222 BP.
 XX
 AC ADQ24610;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7430.
 XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
 KW ds.
 XX Homo sapiens.
 OS
 XX WO2004048938-A2.
 PN 10-JUN-2004.
 PD 26-NOV-2003; 2003WO-US038193.
 XX 26-NOV-2002; 2002US-0429739P.
 PR (PROT-) PROTEIN DESIGN LABS INC.
 XX


```
PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 7430; 210pp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytosstatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 4222 BP; 672 A; 1411 C; 1348 G; 725 T; 0 U; 66 Other;
SQ
Query Match 90.6%; Score 15.4; DB 12; Length 4222;
Best Local Similarity 94.1%; Pred. No. 4.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
Db 3295 GCCCAGCGTTGGCCAAAG 3311
RESULT 48
ADRO7536
ID ADRO7536 standard; cDNA; 4232 BP.
XX
XX ADRO7536;
AC
XX
XX 04-NOV-2004 (first entry)
DT
DE
DE Full length human cDNA useful for treating neurological disease Seq 1042.
XX
XX gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
XX osteoporosis; neurological disease; Alzheimer's disease;
XX Parkinson's disease; dementia; short memory; cancer;
XX sense or motor function; emotional reaction; fear response; panic;
XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
XX tranquiliser.
XX
XX Homo sapiens.
XX
XX EP147413-A2.
XX
XX 18-AUG-2004.
XX
XX 12-FEB-2004; 2004EP-00003145.
XX
XX 14-FEB-2003; 2003JP-00102207.
XX
XX 09-MAY-2003; 2003JP-00131452.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI; 2004-583265/57.
XX
XX P-PSDB; ADRO9492.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PI
```

```
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 1042; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytosstatic and tranquiliser activities. This polynucleotide is a full
CC length human cDNA sequence of the invention. NOTE: This sequence is not
CC given in the sequence listing of the specification but can be obtained on
CC CD-ROM from the European Patent Office, Vienna Sub-office.
XX
XX Sequence 4232 BP; 665 A; 1460 C; 1371 G; 736 T; 0 U; 0 Other;
SQ
Query Match 90.6%; Score 15.4; DB 13; Length 4232;
Best Local Similarity 94.1%; Pred. No. 4.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
Db 3323 GCCCAGCGTTGGCCAAG 3339
RESULT 49
ABA21144/C
ID ABA21144 standard; DNA; 9699 BP.
XX
XX ABA21144;
AC
XX
XX 23-JAN-2002 (first entry)
DT
DE
DE Human nervous system related polynucleotide SEQ ID NO 13475.
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
XX antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
XX Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
```



```
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9699 BP; 1831 A; 2869 C; 3150 G; 1849 T; 0 U; 0 Other;
Query Match 90.6%; Score 15.4; DB 5; Length 9699;
Best Local Similarity 94.1%; Pred. No. 4.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCGGAG 17
Db 678 GCCCAGCGTTGGCGGAG 662
RESULT 50
ABA21143/C
ID ABA21143 standard; DNA; 17457 BP.
XX
AC ABA21143;
XX
XX 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 13474.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antieckling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotrophic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233085P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
```



```

PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234283P.
PR 21-SEP-2000; 2000US-0234284P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX Disclosure; SEQ ID NO 13473; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
XX (AB114678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 17458 BP; 3531 A; 4969 C; 5361 G; 3597 T; 0 U; 0 Other;
XX
XX Query Match 90.6%; Score 15.4; DB 5; Length 17458;
XX Best Local Similarity 94.1%; Pred. No. 4.6e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GCCCAGCTTGGCGGAG 17
XX |||||
XX Db 8434 GCCCAGCTTGGCGGAG 8418
XX
XX RESULT 52
XX ABD33426
XX ID ABD33426 standard; DNA; 63761 BP.
XX AC ABD33426;
XX
XX DT 18-NOV-2004 (first entry)
XX DE Murine cancer-associated (CA) gene MD07-080.
XX
XX KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;

```

KW ds; cancer; cytostatic.

XX OS Mus musculus.

XX PN WO2004058146-A2.

XX XX 15-JUL-2004.

XX PF 15-DEC-2003; 2003WO-US040081.

XX PR 17-DEC-2002; 2002US-00322281.

XX XX (SAGR-) SAGRES DISCOVERY INC.

XX PA Morris DW, Malandro MS;

XX DR WPI; 2004-499109/47.

XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.

XX PS Disclosure; SEQ ID NO 557; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-associated (CA) nucleic acids encoding them. The invention also relates to a method for treating cancers involving administering to a patient an inhibitor of CAP, and a method of screening for anticancer activity in a potential drug involving providing a cell that expresses a CA gene, contacting a tissue sample derived from a cancer cell with an anticancer drug candidate and monitoring the effect of the anticancer drug candidate on expression of the CA gene. The CAP proteins are useful for detecting cancer associated with expression of a CAP protein in a test cell sample and for screening for a bioactive agent capable of modulating the activity of a CAP protein. The CA nucleic acids are useful for diagnosing cancer, involving determining the expression of a CA nucleic acid in a tissue. This sequence represents a murine CA gene of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 63761 BP; 13712 A; 15513 C; 16452 G; 15636 T; 0 U; 2448 Other;

Query Match 90.6%; Score 15.4; DB 13; Length 63761;

Best Local Similarity 94.1%; Pred. No. 4.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17

|||||

Db 34522 GCCCAGCGTTGGCCGAG 34538

RESULT 53

ABD33428

ID ABD33428 standard; DNA; 68732 BP.

XX AC ABD33428;

XX DT 18-NOV-2004 (first entry)

XX DE Human cancer-associated (CA) gene HD07-080.

XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KW ds; cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO2004058146-A2.

XX XX 15-JUL-2004.

XX PF 15-DEC-2003; 2003WO-US040081.

XX PR 17-DEC-2002; 2002US-00322281.

XX (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2004-499109/47.

XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.

XX PS Claim 16; SEQ ID NO 560; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-associated (CA) nucleic acids encoding them. The invention also relates to a method for treating cancers involving administering to a patient an inhibitor of CAP, and a method of screening for anticancer activity in a potential drug involving providing a cell that expresses a CA gene, contacting a tissue sample derived from a cancer cell with an anticancer drug candidate and monitoring the effect of the anticancer drug candidate on expression of the CA gene. The CAP proteins are useful for detecting cancer associated with expression of a CAP protein in a test cell sample and for screening for a bioactive agent capable of modulating the activity of a CAP protein. The CA nucleic acids are useful for diagnosing cancer, involving determining the expression of a CA nucleic acid in a tissue. This sequence represents a human CA gene of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 68732 BP; 14866 A; 17372 C; 18480 G; 16597 T; 0 U; 1417 Other;

Query Match 90.6%; Score 15.4; DB 13; Length 68732;

Best Local Similarity 94.1%; Pred. No. 4.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17

|||||

Db 38569 GCCCAGCGTTGGCCGAG 38585

RESULT 54

ABQ67195_4

Continuation (5 of 5) of ABQ67195 from base 400001 (Listeria innocua contig DNA sequence
WP Sequence split into 5 fragments LOCUS ABQ67195 Accession ABQ67195

WP	Fragment Name	Begin	End
WP	ABQ67195_0	1	110000
WP	ABQ67195_1	100001	210000
WP	ABQ67195_2	200001	310000
WP	ABQ67195_3	300001	410000
WP	ABQ67195_4	400001	495269

Query Match 90.6%; Score 15.4; DB 6; Length 95269;

Best Local Similarity 94.1%; Pred. No. 4.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17

|||||

Db 813 GCCCAGCGTTGGCTGAG 829

RESULT 55

ABQ69245_13/c

Continuation (14 of 31) of ABQ69245 from base 1300001 (Listeria innocua DNA sequence #68
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession ABQ69245

WP	Fragment Name	Begin	End
WP	ABQ69245_00	1	110000
WP	ABQ69245_01	100001	210000
WP	ABQ69245_02	200001	310000
WP	ABQ69245_03	300001	410000
WP	ABQ69245_04	400001	510000
WP	ABQ69245_05	500001	610000
WP	ABQ69245_06	600001	710000
WP	ABQ69245_07	700001	810000

```
WP ABQ69245_08 800001 910000
WP ABQ69245_09 900001 1010000
WP ABQ69245_10 100001 1110000
WP ABQ69245_11 1100001 1210000
WP ABQ69245_12 1200001 1310000
WP ABQ69245_13 1300001 1410000
WP ABQ69245_14 1400001 1510000
WP ABQ69245_15 1500001 1610000
WP ABQ69245_16 1600001 1710000
WP ABQ69245_17 1700001 1810000
WP ABQ69245_18 1800001 1910000
WP ABQ69245_19 1900001 2010000
WP ABQ69245_20 2000001 2110000
WP ABQ69245_21 2100001 2210000
WP ABQ69245_22 2200001 2310000
WP ABQ69245_23 2300001 2410000
WP ABQ69245_24 2400001 2510000
WP ABQ69245_25 2500001 2610000
WP ABQ69245_26 2600001 2710000
WP ABQ69245_27 2700001 2810000
WP ABQ69245_28 2800001 2910000
WP ABQ69245_29 2900001 3010000
WP ABQ69245_30 3000001 3011208
```

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 6; Length 110000;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 GCCCAGCGTTGGCGGAG 17
    |||||
Db 58141 GCCCAGCGTTGGCTGAG 58125
```

RESULT 56

ABQ67195_3
Continuation (4 of 5) of ABQ67195 from base 300001 (Listeria innocua contig DNA sequence
WP Sequence split into 5 fragments LOCUS ABQ67195 Accession Abq67195
WP Fragment Name Begin End
WP ABQ67195_0 1 110000
WP ABQ67195_1 100001 210000
WP ABQ67195_2 200001 310000
WP ABQ67195_3 300001 410000
WP ABQ67195_4 400001 495269

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 6; Length 110000;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 GCCCAGCGTTGGCGGAG 17
    |||||
Db 100813 GCCCAGCGTTGGCTGAG 100829
```

RESULT 57

ABA03041_13/c
Continuation (14 of 30) of ABA03041 from base 1300001 (Listeria monocytogenes EGD-e gene
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
WP Fragment Name Begin End
WP ABA03041_00 1 110000
WP ABA03041_01 100001 210000
WP ABA03041_02 200001 310000
WP ABA03041_03 300001 410000
WP ABA03041_04 400001 510000
WP ABA03041_05 500001 610000
WP ABA03041_06 600001 710000
WP ABA03041_07 700001 810000
WP ABA03041_08 800001 910000
WP ABA03041_09 900001 1010000
WP ABA03041_10 1000001 1110000
WP ABA03041_11 1100001 1210000
WP ABA03041_12 1200001 1310000
WP ABA03041_13 1300001 1410000
WP ABA03041_14 1400001 1510000

```
WP ABA03041_15 1500001 1610000
WP ABA03041_16 1600001 1710000
WP ABA03041_17 1700001 1810000
WP ABA03041_18 1800001 1910000
WP ABA03041_19 1900001 2010000
WP ABA03041_20 2000001 2110000
WP ABA03041_21 2100001 2210000
WP ABA03041_22 2200001 2310000
WP ABA03041_23 2300001 2410000
WP ABA03041_24 2400001 2510000
WP ABA03041_25 2500001 2610000
WP ABA03041_26 2600001 2710000
WP ABA03041_27 2700001 2810000
WP ABA03041_28 2800001 2910000
WP ABA03041_29 2900001 2944528
```

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 6; Length 110000;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 GCCCAGCGTTGGCGGAG 17
    |||||
Db 51366 GCCCAGCGTTGGCTGAG 51350
```

RESULT 58

ABQ81846/c
ID ABQ81846 standard; DNA; 349980 BP.

XX AC ABQ81846;

XX DT 19-NOV-2002 (first entry)

XX DE Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1102.

XX KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
XX KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
XX KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
XX KW rotavirus; food composition; pharmaceutical composition; gene; ds.

XX OS Bifidobacterium longum.

XX OS Synthetic.

XX EN EPI227152-Al.

XX PD 31-JUL-2002.

XX PF 30-JAN-2001; 2001EP-00102050.

XX PR 30-JAN-2001; 2001EP-00102050.

XX PA (NEST) SOC PROD NESTLE SA.

XX DR WPI; 2002-668397/72.

XX PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
a probe or primer for detecting and/or identifying Bifidobacterium longum
in a biological sample.

XX PS Disclosure; SEQ ID NO 1102; 80pp; English.

XX CC The present invention describes a polynucleotide (I) comprising a
sequence of a Bifidobacterium genome selected from the nucleotide
sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
least 90% identity or which hybridises with the sequences given in
ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
fusion protein, comprising a sequence selected from 1097 sequences given
in ABP65258 to ABP63354 ligated in frame to a polynucleotide encoding a
heterologous polypeptide. (I) has antidiarrheic and antibacterial
activities, and can be used as an inhibitor of Salmonella. (I) (which is
a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be

used for preventing and/or treating diarrhoea brought about by pathogenic bacteria and/or rotavirus. The carrier is a food composition selected from milk, yogurt, curd, cheese, fermented milks, milk based products, ice-creams, fermented cereal based products, milk based powders, infant formula, pet food or a pharmaceutical composition selected from tablets, liquid bacterial suspensions, dried oral supplement, wet oral supplement, dry tube feeding or wet tube feeding. (I) is useful in DNA arrays or chips to carry out analysis of the expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent Bifidobacterium related nucleotide sequences given in the Sequence Listing from the present invention but not mentioned further within the specification. N.B. The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Sequence 349980 BP; 69200 A; 103414 C; 105147 G; 72219 T; 0 U; 0 Other;

Query Match 90.6%; Score 15.4; DB 6; Length 349980;

Best Local Similarity 94.1%; Pred. No. 4.8e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAG 17

DB 322894 GCCCAGCTTGCCGAG 322878

RESULT 59

ABQ81847/c

ID ABQ81847 standard; DNA; 349980 BP.

AC ABQ81847;

DT 19-NOV-2002 (first entry)

DE Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1103.

Human Bifidobacterium longum NCC2705; Bifidobacterium; bacterial; antidiarrheic; antibacterial; inhibitor of Salmonella; detection; identification; lactic acid bacterium; diarrhoea; pathogenic bacteria; rotavirus; food composition; pharmaceutical composition; gene; ds.

Bifidobacterium longum.

Synthetic.

EP1227152-A1.

31-JUL-2002.

30-JAN-2001; 2001EP-00102050.

30-JAN-2001; 2001EP-00102050.

(NEST) SOC PROD NESTLE SA.

WPI; 2002-668397/72.

Novel polynucleotide comprising Bifidobacterium genome sequence useful as a probe or primer for detecting and/or identifying Bifidobacterium longum in a biological sample.

Disclosure; SEQ ID NO 1103; 80pp; English.

The present invention describes a polynucleotide (I) comprising a sequence of a Bifidobacterium genome selected from the nucleotide sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at least 90% identity or which hybridises with the sequences given in ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a fusion protein, comprising a sequence selected from 1097 sequences given in ABP65258 to ABP6354 ligated in frame to a polynucleotide encoding a heterologous polypeptide. (I) has antidiarrheic and antibacterial activities, and can be used as an inhibitor of Salmonella. (I) (which is a probe) is useful for the detection and/or identification of Bifidobacterium longum in a biological sample. A carrier containing the

lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM 1-2618) can be used for preventing and/or treating diarrhoea brought about by pathogenic bacteria and/or rotavirus. The carrier is a food composition selected from milk, yogurt, curd, cheese, fermented milks, milk based products, ice-creams, fermented cereal based products, milk based powders, infant formula, pet food or a pharmaceutical composition selected from tablets, liquid bacterial suspensions, dried oral supplement, wet oral supplement, dry tube feeding or wet tube feeding. (I) is useful in DNA arrays or chips to carry out analysis of the expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent Bifidobacterium related nucleotide sequences given in the Sequence Listing from the present invention but not mentioned further within the specification. N.B. The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Sequence 349980 BP; 69124 A; 105769 C; 105545 G; 69542 T; 0 U; 0 Other;

Query Match 90.6%; Score 15.4; DB 6; Length 349980;

Best Local Similarity 94.1%; Pred. No. 4.8e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAG 17

DB 22894 GCCCAGCTTGCCGAG 22878

RESULT 60

ABV57592/c

ID ABV57592 standard; cDNA; 337 BP.

AC ABV57592;

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 57583.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 11075; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 337 BP; 88 A; 101 C; 60 G; 88 T; 0 U; 0 Other;
 Query Match 88.2%; Score 15; DB 5; Length 337;
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 CCACGCTTGGCCGAG 17
 Db 216 CCACGCTTGGCCGAG 202
 |||||
 RESULT 61
 ADT19568/c
 ID ADT19568 standard; cDNA; 3303 BP.
 AC ADT19568;
 XX
 DT 13-JAN-2005 (first entry)
 XX
 DE Plant cDNA, Seq ID 4894.
 XX
 KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
 KW drought tolerance; disease resistance; galactomannan production;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW lignin production; extreme osmotic condition tolerance;
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KW seed protein yield.
 XX
 OS Viridiplantae.
 XX
 PN US2004216190-A1.
 XX
 PD 28-OCT-2004.
 XX
 PF 18-DEC-2003; 2003US-00739930.
 XX
 PR 28-APR-2003; 2003US-00424599.
 PR 28-APR-2003; 2003US-00425115.
 XX
 PA (KOVA/) KOVALIC D K.
 XX
 PI Kovalic DK;
 XX
 DR WPI; 2004-757369/74.
 XX
 PT New recombinant DNA constructs useful in the field of biochemistry and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX
 PS Claim 1; SEQ ID NO 4894; 14pp; English.
 XX
 CC The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed
 CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for

CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biochemistry and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20040216190.
 XX
 SQ Sequence 3303 BP; 637 A; 1063 C; 991 G; 612 T; 0 U; 0 Other;
 Query Match 88.2%; Score 15; DB 13; Length 3303;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 CCCACGCTTGGCCGA 16
 Db 619 CCCACGCTTGGCCGA 605
 |||||
 RESULT 62
 AAS37371/c
 ID AAS37371 standard; cDNA; 344 BP.
 XX
 AC AAS37371;
 XX
 DT 17-DEC-2001 (first entry)
 XX
 DE Novel human diagnostic and therapeutic gene #429.
 XX
 KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200166753-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 09-MAR-2001; 2001WO-US007787.
 XX
 PR 09-MAR-2000; 2000US-0188609P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
 PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
 DR WPI; 2001-530177/58.
 XX
 PT New polynucleotides and polypeptides, useful for diagnosis and treatment
 PT of breast, lung and colon cancer.
 XX
 PS Claim 1; Page 704; 1193pp; English.
 XX
 CC The invention relates to new polynucleotides and polypeptides, useful for
 CC diagnosis and treatment of breast, lung and colon cancer. The sequences

CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention

XX SQ Sequence 344 BP; 146 A; 53 C; 85 G; 59 T; 0 U; 1 Other;

Query Match 84.7%; Score 14.4; DB 4; Length 344;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CCCACGCTTGGCCGAG 17

DB 252 CCCACGCTTGGCCGTAG 237

RESULT 63

AAH09789/C

ID AAH09788 standard; cDNA; 546 BP.

XX AC

XX AAH09788;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (3'-primer) SEQ ID NO:6623.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.

XX XN EPI074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX XN (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.

XX Claim 3; SEQ ID NO 6623; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX SQ Sequence 546 BP; 119 A; 130 C; 124 G; 165 T; 0 U; 8 Other;

Query Match 84.7%; Score 14.4; DB 4; Length 546;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCCCAGCTTGGCCGA 16

DB 505 GCCCAGCTTGGCCGA 490

RESULT 64

ABV54022

ID ABV54022 standard; cDNA; 570 BP.

XX AC

XX ABV54022;

XX DT 17-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 54013.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.

XX XN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX XN (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 10452; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

```
SQ Sequence 570 BP; 119 A; 167 C; 152 G; 129 T; 0 U; 3 Other;
Query Match      84.7%; Score 14.4; DB 5; Length 570;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
   |||||
Db 140 GCCCAGCGTTGACCGA 155

RESULT 65
AAH08354
ID AAH08354 standard; cDNA; 680 BP.
XX
AC AAH08354;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:5189.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
FN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUN-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 1; SEQ ID NO 5189; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

SQ Sequence 680 BP; 143 A; 194 C; 151 G; 184 T; 0 U; 8 Other;
Query Match      84.7%; Score 14.4; DB 4; Length 680;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
   |||||
Db 633 GCCCAGCGTTGGCCCA 648

RESULT 66
AAL55183/c
ID AAL55183 standard; DNA; 992 BP.
XX
AC AAL55183;
XX
DT 01-MAY-2003 (first entry)
XX
DE Verticillium dahliae defoliating plant DNA, SEQ ID NO 2.
XX
KW Defoliating; non-defoliating; pathogen; Verticillium dahliae; nested PCR; olive; cotton; epidemiology; plant; ds.
XX
OS Verticillium dahliae.
XX
FN WO200297131-A1.
XX
PD 05-DEC-2002.
XX
PF 29-MAY-2002; 2002WO-ES000258.
PR 29-MAY-2001; 2001ES-00001240.
XX
PA (NEWB-) NEWBIOTECHNIC SA.
XX
PA (UYCO-) UNIV CORDOBA.
XX
XX Mercado Blanco J, Rodriguez Jurado D, Perez Artes E;
PI Jimenez Diaz RM;
XX
DR WPI; 2003-140494/13.
XX
XX Identifying D and ND pathotypes of Verticillium dahliae, useful for early diagnosis of infection in olives and cotton, by nested polymerase chain reaction.
XX
PS Disclosure; Fig 2; 68pp; Spanish.
XX
XX The invention relates to a novel method for identifying defoliating and/or non-defoliating pathogenic types of Verticillium dahliae in a sample of plant material by use of a nested polymerase chain reaction. The method is used to differentiate between the pathotypes of V. dahliae of olives and cotton, e.g. to establish certification of disease-free plants, and also to study the epidemiology and control of V. dahliae. This polynucleotide sequence represents the DNA of a Verticillium dahliae defoliating plant of the invention
XX
SQ Sequence 992 BP; 212 A; 229 C; 306 G; 245 T; 0 U; 0 Other;
Query Match      84.7%; Score 14.4; DB 8; Length 992;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
   |||||
Db 432 GCCCAGCGTTGGCCGA 417

RESULT 67
AAV63901/c
ID AAV63901 standard; DNA; 1520 BP.
XX
```

AC AAV63901;
 DT 20-JAN-1999 (first entry)
 XX Helicobacter pylori DNA insert full restriction map of plasmid pILL405.
 DE Helicobacter pylori; amIE; aliphatic amidase; gastritis; dyspepsia;
 XX chronic gastroduodenal disorder; peptic ulcer; porcine gastric ulcer;
 KW detection; infection; ss.
 KW Helicobacter pylori.
 OS WO9844094-A2.
 XX 08-OCT-1998.
 XX 27-MAR-1999; 98WO-EP001824.
 PF 28-MAR-1997; 97US-0041745P.
 PR (INSP) INST PASTEUR.
 XX (INRM) INST NAT SANTE & RECH MEDICALE.
 PA De Reuse H, Skouloubris S, Labigne A;
 PI WPI; 1998-557106/47.
 DR New Helicobacter aliphatic amidase AmIE polypeptides and their encoding
 PT sequences - used in diagnosis, treatment and prevention of Helicobacter
 PT sp. infections in humans and animals.
 XX Claim 2; Fig 4b; 51pp; English.
 PS The present sequence represents the Helicobacter pylori DNA insert full
 CC restriction map of plasmid pILL405. The present invention describes H.
 CC pylori aliphatic amidase AmIE. AmIE polypeptides and antibodies form
 CC immunogenic, pharmaceutical and therapeutic compositions which can be
 CC used in methods for detecting, treating or preventing Helicobacter sp.
 CC (particularly H. pylori and H. heilmannii) infections, especially chronic
 CC gastroduodenal disorders like gastritis, dyspepsia and peptic ulcers in
 CC man, and porcine gastric ulcers in pigs. They are administered in
 CC immunologically/pharmacologically effective amounts by an oral,
 CC intradermal, intramuscular, intravenous or mucosal route to a patient. A
 CC detection kit for a Helicobacter infection comprises any AmIE
 CC polynucleotides with any of the vectors given in the present invention.
 CC The polypeptides are used to in a process to degrade acrylamide,
 CC acetamide, propionamide and isobutyramide. The polypeptides are also used
 CC to screen for active substrates that inhibit Helicobacter sp. amidase
 CC activity. The polynucleotides encoding the aliphatic amidase AmIE
 CC polypeptides are the first to be characterized at the molecular level.
 CC Methods involving such polypeptides are preferred to urease-based methods
 CC because of the presence of urea positive bacteria in porcine
 CC gastrointestinal tracts
 XX Sequence 1520 BP; 461 A; 261 C; 385 G; 413 T; 0 U; 0 Other;
 SQ Query Match 84.7%; Score 14.4; DB 2; Length 1520;
 Best Local Similarity 93.8%; Pred. No. 1.4e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 CCCAGCTTGGCCGAG 17
 |||||
 Db 1454 CCCAGCTTGGCCGAG 1439
 RESULT 68
 ACD05884/c
 ID ACD05884 standard; cDNA; 1531 BP.
 XX AC ACD05884;
 XX 06-AUG-2003 (first entry)
 DT

DE Novel human contig #58.
 XX Human; angiogenesis; cytokine; cell proliferation; pluripotent;
 KW cell differentiation; totipotent; stem cell; transplantation; bio-sensor;
 KW neuroepithelial cell; autoimmune disease; neural cell; genetic disorder;
 KW nerve; brain tissue; central nervous system disease;
 KW peripheral nervous system disease; neuropathy; haematopoiesis; bone;
 KW myeloid disorder; lymphoid cell disorder; platelet disorder; tendon;
 KW regeneration; cartilage; tendon; ligament; nerve tissue growth;
 KW tissue repair; wound healing; burn; ulcer; osteoporosis; cancer;
 KW osteoarthritis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung fibrosis; liver fibrosis; reperfusion injury;
 KW immune deficiency; infection; autoimmune disorder; allergic reaction;
 KW thrombolysis; thrombosis; coagulation disorder; hereditary disorder;
 KW biorhythm; circadian cycle; fertility; metabolism; catabolism; anabolism;
 KW neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;
 KW haemostatic; vulnery; antiulcer; osteopathic; antiarthritic;
 KW vasotropic; immunostimulant; antibacterial; fungicide; immunosuppressive;
 KW antirheumatic; antidiabetic; antiasthmatic; cytostatic; virucide;
 KW expressed sequence tag; EST; ss.
 XX Homo sapiens.
 OS WO2003023013-A2.
 XX 20-MAR-2003.
 PD 13-SEP-2002; 2002WO-US029001.
 PF 13-SEP-2001; 2001US-0322511P.
 PR 12-SEP-2002; 2002US-00243552.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Yang Y, Wang Z, Wang G, Ma Y;
 XX WPI; 2003-313249/30.
 PI P-PSDB; ABO00807.
 DR Novel nucleic acids and polypeptides for diagnosis, treatment of central
 XX and peripheral nervous system diseases and neuropathies, such as
 PT Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 PT lateral sclerosis.
 XX Example 2; SEQ ID NO 730; 300pp; English.
 PS The present invention relates to the isolation of novel human
 CC polynucleotide sequences and their encoding polypeptides. The novel
 CC polypeptides exhibit activities relating to angiogenesis, cytokine, cell
 CC proliferation, cell differentiation, antiinflammatory, and stem cell
 CC growth factor activities. The polypeptides are involved in the
 CC proliferation, differentiation and survival of pluripotent and totipotent
 CC stem cells, and are useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals and
 CC development of bio-sensors. The polypeptides can be used to manipulate
 CC stem cells in culture to give rise to neuroepithelial cells that can be
 CC used to augment or replace cells damaged by illness, autoimmune disease,
 CC accidental damage or genetic disorders. The polypeptides induce the
 CC proliferation of neural cells and regeneration of nerve and brain tissue
 CC and are useful for the treatment of central and peripheral nervous system
 CC diseases and neuropathies, such as Alzheimer's, Parkinson's disease,
 CC Huntington's disease, amyotrophic lateral sclerosis (ALS). The
 CC polypeptides are also involved in chemotactic or chemokinetic activity,
 CC regulation of haematopoiesis and are useful for treating myeloid or
 CC lymphoid cell disorders, platelet disorders such as thrombocytopaenia and
 CC for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
 CC growth. In tissue repair, healing of burns, incisions, ulcers, for
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders, and
 CC periodontal disease. The polypeptides are also useful for gut protection
 CC or regeneration and treatment of lung or liver fibrosis, reperfusion
 CC injury in various tissues, various immune deficiencies and disorders
 CC including severe combined immunodeficiency (SCID), bacterial or fungal
 CC infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid

CC arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and
 CC conditions, such as asthma or other respiratory problems. The
 CC polypeptides are involved in thrombolysis or thrombosis and are useful in
 CC treatment of various coagulation disorders (including hereditary
 CC disorders such as haemophilia) or to enhance coagulation and other
 CC haemostatic events in treating wounds resulting from trauma, surgery or
 CC other causes. The polypeptides exhibit immune stimulating or immune
 CC suppressing activity, and are useful for treating autoimmune diseases or
 CC cancer. They also inhibit the growth, infection or function of infectious
 CC agents such as bacteria, fungi, viruses, effect biorhythms or circadian
 CC cycles of rhythms, fertility of male or female subjects, metabolism,
 CC catabolism, and anabolism. ACD05827-ACD06027 represent novel contigs
 CC assembled using expressed sequence tag (EST) sequences as seeds. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1531 BP; 407 A; 409 C; 451 G; 264 T; 0 U; 0 Other;

Query Match 84.7%; Score 14.4; DB 8; Length 1531;
 Best Local Similarity 93.8%; Pred. No. 1.4e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGA 16
 Db 120 GCCCATGCTTGGCCGA 105
 ||||| ||||| |||||

RESULT 69
 AAV63900/C
 ID AAV63900 standard; DNA; 1675 BP.

XX AAV63900;

XX 20-JAN-1999 (first entry)

XX Helicobacter pylori DNA insert sequence of plasmid pILL405.

XX Helicobacter pylori; amIE; aliphatic amidase; gastritis; dyspepsia;
 KW chronic gastroduodenal disorder; peptic ulcer; porcine gastric ulcer;
 KW detection; infection; ss.

XX Helicobacter pylori.

XX WO9844094-A2.

XX 08-OCT-1998.

XX 27-MAR-1998; 98WO-EP001824.

XX 28-MAR-1997; 97US-0041745P.

XX (INSP) INST PASTEUR.

XX (INRM) INST NAT SANTE & RECH MEDICALE.

XX De Reuse H, Skouloubris S, Labigne A;

XX WPI; 1998-557106/47.

XX New Helicobacter aliphatic amidase AmIE polypeptides and their encoding
 PT sequences - used in diagnosis, treatment and prevention of Helicobacter
 PT sp. infections in humans and animals.

XX Claim 2; Fig 4a; 51pp; English.

XX The present sequence represents the Helicobacter pylori DNA insert
 CC sequence of plasmid pILL405. The present invention describes H. pylori
 CC aliphatic amidase AmIE. AmIE polypeptides and antibodies form
 CC immunogenic, pharmaceutical and therapeutic compositions which can be
 CC used in methods for detecting, treating or preventing Helicobacter sp.
 CC (particularly H. pylori and H. heilmannii) infections, especially chronic
 CC gastroduodenal disorders like gastritis, dyspepsia and peptic ulcers in
 CC man, and porcine gastric ulcers in pigs. They are administered in

CC immunologically/pharmaceutically effective amounts by an oral,
 CC intradermal, intramuscular, intravenous or mucosal route to a patient. A
 CC detection kit for a Helicobacter infection comprises any AmIE
 CC polynucleotides with any of the vectors given in the present invention.
 CC The polypeptides are used in a process to degrade acrylamide,
 CC acetamide, propionamide and isobutyramide. The polypeptides are also used
 CC to screen for active substrates that inhibit Helicobacter sp. amidase
 CC activity. The polynucleotides encoding the aliphatic amidase AmIE
 CC polypeptides are the first to be characterized at the molecular level.
 CC Methods involving such polypeptides are preferred to urase-based methods
 CC because of the presence of urea positive bacteria in porcine
 CC gastrointestinal tracts
 XX

SQ Sequence 1675 BP; 503 A; 298 C; 424 G; 450 T; 0 U; 0 Other;

Query Match 84.7%; Score 14.4; DB 2; Length 1675;

Best Local Similarity 93.8%; Pred. No. 1.4e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAG 17

Db 1534 CCCACGCTTGGCCGAG 1519
 ||||| ||||| |||||

RESULT 70

ABZ35358/c

ID ABZ35358 standard; cDNA; 1889 BP.

XX ABZ35358;

XX 05-FEB-2003 (first entry)

XX Human gene expression profile polynucleotide SEQ ID NO 469.

XX Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.

XX Homo sapiens.

XX WO200274979-A2.

XX 26-SEP-2002.

XX 20-MAR-2002; 2002WO-US008456.

XX 20-MAR-2001; 2001US-0276947P.

XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.

XX Wan J, Wang Y;

XX WPI; 2002-740862/80.

XX New gene expression profile generated from primary, endothelial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer.

XX Disclosure; Page 617-618; 850pp; English.

XX The invention relates to a gene expression profile comprising one or more
 CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,

CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes, DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC antifungal agents
CC

SQ Sequence 1889 BP; 476 A; 464 C; 494 G; 455 T; 0 U; 0 Other;

Query Match 84.7%; Score 14.4; DB 6; Length 1889;
Best Local Similarity 93.8%; Pred. NO. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGGCCGAG 17
D5 18 CCCACGCTTGGCCGAG 3

RESULT 71

ACN40151/c
ID ACN40151 standard; cDNA; 3681 BP.

XX AC ACN40151;

XX 18-NOV-2004 (first entry)

XX Tumour-associated antigenic target (TAT) cDNA DNA326301, SEQ ID NO:4752.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

XX 02-OCT-2002; 2002US-0414971P.

XX (GETH) GENENTECH INC.

XX Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

XX P-PSDB; ABM81845.

XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

XX Claim 1; SEQ ID NO 4752; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic

CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention

SQ Sequence 3681 BP; 1095 A; 903 C; 968 G; 715 T; 0 U; 0 Other;

Query Match 84.7%; Score 14.4; DB 13; Length 3681;
Best Local Similarity 93.8%; Pred. NO. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGA 16
D5 149 GCCCAGCTTGGCCGA 134

RESULT 72

ABK89239/c
ID ABK89239 standard; DNA; 13608 BP.

XX ABK89239;

XX 21-OCT-2002 (first entry)

XX Human Na/Cl-dependent organic solute transporter gene.

XX Human; ds; gene; Na/Cl-dependent organic solute cotransporter;
KW chromosome 5; neuron tumour; single nucleotide polymorphism; SNP.

XX Homo sapiens.

XX Key Location/Qualifiers
FT variation replace(26,C)

FT /*tag= a

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(36,R)

FT /*tag= b

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(131,R)

FT /*tag= c

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(1330,C)

FT /*tag= d

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(1603,C)

FT /*tag= e

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(1773,T)

FT /*tag= f

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(1885,T)

FT /*tag= g

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(1897,T)

FT /*tag= h

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(2010,G)

FT /*tag= i

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(2260,G)

FT /*tag= j

FT /*standard_name= "Single nucleotide polymorphism"

FT variation replace(2314,T)

FT /*tag= k

FT /*standard_name= "Single nucleotide polymorphism"

FT /standard_name="Single nucleotide polymorphism"
FT replace(2650,C)
FT /*tag= l
FT /standard_name="Single nucleotide polymorphism"
FT replace(4944,A)
FT /*tag= m
FT /standard_name="Single nucleotide polymorphism"
FT replace(5230,T)
FT /*tag= n
FT /standard_name="Single nucleotide polymorphism"
FT replace(5481,G)
FT /*tag= o
FT /standard_name="Single nucleotide polymorphism"
FT replace(5765,T)
FT /*tag= p
FT /standard_name="Single nucleotide polymorphism"
FT replace(6097,A)
FT /*tag= q
FT /standard_name="Single nucleotide polymorphism"
FT replace(7194,V)
FT /*tag= r
FT /standard_name="Single nucleotide polymorphism"
FT replace(1117,G)
FT /*tag= s
FT /standard_name="Single nucleotide polymorphism"
FT replace(11398,T)
FT /*tag= t
FT /standard_name="Single nucleotide polymorphism"
FT replace(12464,A)
FT /*tag= u
FT /standard_name="Single nucleotide polymorphism"
FT replace(12736,C)
FT /*tag= v
FT /standard_name="Single nucleotide polymorphism"

XX WO200253741-A2.

XX 11-JUL-2002.

XX 02-JAN-2002; 2002WO-US000111.

XX 03-JAN-2001; 2001US-00752821.

PR 22-MAY-2001; 2001US-00861846.

XX (PEKE) PE CORP NY.

XX Gugler K, Brandon RC, Di Francesco V, Beasley EM;

XX WPT; 2002-583618/62.

XX Novel isolated human peptide useful for treating disorders characterized by the absence of, inappropriate or unwanted expression of the receptor protein, and as immunogens to raise antibodies.

PS Claim 4; Fig 3; 88pp; English.

XX The invention relates to an isolated human peptide (S1) comprising an Na/C1-dependent organic solute cotransporter family member, an allelic variant or orthologue of (S1) encoded by a polynucleotide hybridising under stringent conditions to the opposite strand of a sequence appearing as ABK89218 and ABK92239, or a fragment of (S1) with at least 10 contiguous amino acids. A pharmaceutical composition comprising a modulator of S1 is useful for treating a disease or condition mediated by a human protease. S1 and its encoding nucleic acid are useful as models for the development of human therapeutics, for identifying therapeutic proteins, as targets for development of human therapeutic agents. S1 is useful to raise antibodies or to elicit another immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as markers for tissues in which the corresponding protein is preferentially expressed, in drug screening assays, in cell-based or cell-free systems, to identify compounds that modulate cotransporter activity of the protein in its natural state, or an altered form that causes the specific disease or pathology associated

CC with the receptor (e.g. neuron tumour), as bait proteins in a two-hybrid or three-hybrid assay, to provide a target for diagnosing a disease or predisposition to disease mediated by the peptide, in pharmacogenomic analysis, and for treating a disorder characterised by an absence of, in appropriate or unwanted expression of S1. An anti-S1 antibody is useful for isolating and purifying S1, to detect S1 in cells or tissues, in situ, in vitro or in a cell lysate or supernatant, to assess abnormal tissue distribution or abnormal expression during development or progression of a biological condition, and for inhibiting a protein function. S1 nucleic acids are useful as probes, primers, and chemical intermediates, in biological assays, for constructing recombinant vectors, host cells and transgenic animals, for expressing antigenic portions of the S1 proteins, or for designing ribozymes. A vector comprising the S1 nucleic acid is useful for producing an S1 protein or peptide, for conducting cell-based assays involving the protein or its fragment, for identifying protein mutants whose functions are affected, and to produce non-human transgenic animals. The gene for the cotransporter is located on chromosome 5. The present sequence genomic DNA for the cotransporter protein

XX SQ Sequence 13608 BP; 2678 A; 4019 C; 3850 G; 2869 T; 0 U; 192 Other;
Query Match 84.7%; Score 14.4; DB 6; Length 13608;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CCCACGCTTGCCGAG 17

Db 7742 CCCACGCTTGCTGAG 7727

RESULT 73

AED48654
ID AED48654 standard; DNA; 48169 BP.

XX AC AED48654;

XX 15-DEC-2005 (first entry)

DE Streptomyces rimosus paromomycin biosynthetic gene cluster DNA.

XX ds; gene; antibiotic; virucide; aminoglycoside; bacterial infection;
XX viral infection; paromomycin; parE; parS; parC; parM; parT; parU; parQ;
XX parN; parP; parX; parF; parD; parU; parB; parA; parK; parL; parR; parY;
XX parG; parH; P450 heme thiolate protein; cytochrome P450 hydroxylase;
XX non-ribosomal peptide synthetase; polyketide cyclase/reductase;
XX amidinotransferase; DNA-binding protein;
XX 3-amino-2,3-dideoxy-scyllo-inositol-1-dehydrogenase;
XX bifunctional L-glutamine:ketocyclitol aminotransferase I and II;
XX 2-deoxy-scyllo-inosose synthase;
XX UDP-Nac-glucosamine:2-deoxystreptamine-4-glycosyltransferase;
XX ABC transporter, ATP-binding protein; paromomycin 6'-dehydrogenase;
XX Fe-S oxidoreductase; phosphatase; glycosyltransferase;
XX 2-N-acetylhexosamine deacetylase; paromomycin 6'-aminotransferase;
XX extracellular aminoglycoside phosphate phosphatase;
XX aminoglycoside 3'-phosphotransferase; regulatory protein;
XX NDP-hexose synthase; chitinase;
XX nucleoside diphosphate-sugar oxidoreductase;
XX tetR family transcriptional regulator; transcriptional regulator;
XX oxidoreductase.

XX Streptomyces rimosus.

XX WO2005095591-A2.

XX 13-OCT-2005.

XX 15-MAR-2005; 2005WO-DE000499.

XX 01-APR-2004; 2004DE-10017141.

XX (COMB-) COMBINATURE BIOPHARM AG.
XX (UYWU-) UNIV WUPPERTAL.

XX Vente A, Piepersberg W, Wehmeier U, Schmidt-Beissner H;
 PI Abohanab KMA, Welzel K;
 XX WPI; 2005-734187/75.
 DR P-PSDB; AED48614, AED48615, AED48616, AED48617, AED48618, AED48619,
 DR AED48620, AED48621, AED48622, AED48623, AED48624, AED48625, AED48626,
 DR AED48627, AED48628, AED48629, AED48630, AED48631, AED48632, AED48633,
 DR AED48634, AED48635, AED48636, AED48637, AED48638, AED48639, AED48640,
 DR AED48641, AED48642, AED48643, AED48644, AED48645, AED48646, AED48647,
 DR AED48648, AED48649, AED48650, AED48651, AED48652, AED48653,
 DR DBJ; AJ628955.
 XX Isolated proteins or peptides that are expression products of genes
 PT derived from microorganisms producing aminoglycoside antibiotics.
 PT
 XX Example 1; Page 185-199; 367pp; German.
 PS
 XX This invention describes novel isolated proteins or peptides that are
 CC expression products of genes derived from aminoglycoside-producing
 CC microorganisms. The proteins or peptides are functional for a step in the
 CC biosynthesis of aminoglycoside antibiotics, especially butirosin,
 CC gentamicin, neomycin, fortimicin, tobramycin, apramycin, paromomycin,
 CC hygromycin B, ribostamycin, sisomicin, kanamycin, nebramycin,
 CC sel domycin, destomycin, istamycin, sannamycin, dactimicin, sporaricin,
 CC blusomycin, ashimycin, lividomycin and spectinomycin. The invention
 CC also describes a) isolated nucleic acids encoding the proteins or
 CC peptides which are under the control of a regulatory sequence, especially
 CC a promoter; b) transformation vectors containing the nucleic acids; c)
 CC cells containing the nucleic acids and/or the proteins or peptides; d) a
 CC method of producing the cells by transforming cells with vectors as above
 CC ; e) aminoglycosides produced by culturing the cells; f) producing
 CC aminoglycosides by culturing cells as above and isolating aminoglycosides
 CC from the cells and/or culture supernatant and g) pharmaceutical
 CC compositions containing the aminoglycosides. The aminoglycosides are
 CC useful for treating bacterial or viral infections, especially infections
 CC with Gram-positive or -negative bacteria, e.g. *Escherichia coli* or
 CC *Klebsiella*. This sequence encodes proteins (AED48605-AED48653) from
 CC *Streptomyces rimosus* (DBJ reference AJ628955) involved in the synthesis
 CC of paromomycin.
 XX
 SQ Sequence 48169 BP; 6567 A; 17743 C; 17423 G; 6436 T; 0 U; 0 Other;
 Query Match 84.7%; Score 14.4; DB 14; Length 48169;
 Best Local Similarity 93.8%; Pred. No. 1.5e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCCCAGCTTGCCGA 16
 DB 45829 GCCCAGCTTGCTGA 45844
 RESULT 74
 AAH51594/c
 ID AAH51594 standard; DNA; 49312 BP.
 XX
 AC AAH51594;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE Human genomic sequence of the MGST-II gene SEQ ID 485.
 XX Human; biallelic marker; single nucleotide polymorphism; SNP; MGSTII;
 KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;
 KW DMEI; MEI; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;
 KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;
 KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;
 KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotocity;
 KW uridine diphosphate glucuronosyl transferase; UGT2; asthma; hepatotoxicity;
 KW zileuton; ds.
 XX
 OS Homo sapiens.
 XX

PN WO200058508-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-IB0000403.
 XX
 PR 25-MAR-1999; 99US-0126269P.
 PR 30-APR-1999; 99US-0131961P.
 XX
 PA (GEST) GENSET.
 XX
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;
 XX WPI; 2000-638353/61.
 DR
 XX Polynucleotides comprising sequences from malate decarboxylase enzyme-
 PT related biallelic markers used for genotyping.
 PT
 XX Claim 88; Page 656-669; 673pp; English.
 PS
 XX Sequences AAH51110-AAH51593 represent human DNA fragments which contain
 CC biallelic markers. The sequences are related to various human genes
 CC including microsomal glutathione S-transferase II (MGSTII), malate
 CC decarboxylase enzyme (DMEI/MEI), cytochrome P450, glutathione
 CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),
 CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate
 CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH), and uridine
 CC diphosphate glucuronosyl transferases (UGT2). Each of these sequences
 CC contains a biallelic marker/polymorphism, which is represented in the
 CC sequence as a degenerate/undefined base. The genes to which the biallelic
 CC marker containing sequences are related are involved in drug metabolism.
 CC Sequences AAH51594 - AAH51598 represent the genomic sequence of the
 CC MGSTII gene and four alternative MGSTII cDNA sequences. AAB62905-AAB62906
 CC are MGSTII gene products. PCR primers AAH51599 and AAH51600 are used in
 CC an example for the amplification of human genomic DNA fragments. The
 CC invention includes a method of genotyping comprising determining the
 CC identity of a nucleotide at a DME- or MGSTII-related biallelic marker in
 CC a biological sample. The method is used to determine the frequency in
 CC a population of an allele of a DME- or MGST-II related biallelic marker and
 CC to select an individual for inclusion in a clinical trial of a drug
 CC treatment. The method is also used to detect association between allele
 CC and phenotype, and to detect association between haplotype and phenotype.
 CC The polynucleotides are used, in hybridization assays, sequencing assays
 CC or allele specific amplification assays. The method can be used to
 CC determine whether an individual suffers or is at risk of developing
 CC asthma or is at risk of developing hepatotoxicity on treatment with
 CC zileuton
 XX
 SQ Sequence 49312 BP; 13125 A; 10209 C; 11088 G; 14881 T; 0 U; 9 Other;
 Query Match 84.7%; Score 14.4; DB 3; Length 49312;
 Best Local Similarity 93.8%; Pred. No. 1.5e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 CCCACGCTTGCCGAG 17
 DB 12970 CCCACTCTTGCCGAG 12955
 RESULT 75
 ADN46845_20
 Continuation (21 of 21) of ADN46845 from base 2000001 (Thermococcus kodakaraensis KOD1
 WP Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845
 WP Fragment Name Begin End
 WP ADN46845_00 1 110000
 WP ADN46845_01 100001 210000
 WP ADN46845_02 200001 310000
 WP ADN46845_03 300001 410000
 WP ADN46845_04 400001 510000
 WP ADN46845_05 500001 610000
 WP ADN46845_06 600001 710000
 WP ADN46845_07 700001 810000
 WP ADN46845_08 800001 910000

WP ADN46845_09 900001 1010000
WP ADN46845_10 1000001 1110000
WP ADN46845_11 1100001 1210000
WP ADN46845_12 1200001 1310000
WP ADN46845_13 1300001 1410000
WP ADN46845_14 1400001 1510000
WP ADN46845_15 1500001 1610000
WP ADN46845_16 1600001 1710000
WP ADN46845_17 1700001 1810000
WP ADN46845_18 1800001 1910000
WP ADN46845_19 1900001 2010000
WP ADN46845_20 2000001 2089378

Query Match 84.7%; Score 14.4; DB 12; Length 89378;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
|||||

Db 64028 GCCCAGCGTTGGCCGA 64043

RESULT 76

ADN46123_20

Continuation (21 of 21) of ADN46123 from base 2000001 (Thermococcus kodakaraensis KOD1
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123

WP Fragment Name Begin End
WP ADN46123_00 1 110000
WP ADN46123_01 100001 210000
WP ADN46123_02 200001 310000
WP ADN46123_03 300001 410000
WP ADN46123_04 400001 510000
WP ADN46123_05 500001 610000
WP ADN46123_06 600001 710000
WP ADN46123_07 700001 810000
WP ADN46123_08 800001 910000
WP ADN46123_09 900001 1010000
WP ADN46123_10 1000001 1110000
WP ADN46123_11 1100001 1210000
WP ADN46123_12 1200001 1310000
WP ADN46123_13 1300001 1410000
WP ADN46123_14 1400001 1510000
WP ADN46123_15 1500001 1610000
WP ADN46123_16 1600001 1710000
WP ADN46123_17 1700001 1810000
WP ADN46123_18 1800001 1910000
WP ADN46123_19 1900001 2010000
WP ADN46123_20 2000001 2089378

Query Match 84.7%; Score 14.4; DB 12; Length 89378;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
|||||

Db 64028 GCCCAGCGTTGGCCGA 64043

RESULT 77

ADN46464_20

Continuation (21 of 21) of ADN46464 from base 2000001 (Thermococcus kodakaraensis KOD1
WP Sequence split into 21 fragments LOCUS ADN46464 Accession Adn46464

WP Fragment Name Begin End
WP ADN46464_00 1 110000
WP ADN46464_01 100001 210000
WP ADN46464_02 200001 310000
WP ADN46464_03 300001 410000
WP ADN46464_04 400001 510000
WP ADN46464_05 500001 610000
WP ADN46464_06 600001 710000
WP ADN46464_07 700001 810000
WP ADN46464_08 800001 910000
WP ADN46464_09 900001 1010000

WP ADN46464_10 1000001 1110000
WP ADN46464_11 1100001 1210000
WP ADN46464_12 1200001 1310000
WP ADN46464_13 1300001 1410000
WP ADN46464_14 1400001 1510000
WP ADN46464_15 1500001 1610000
WP ADN46464_16 1600001 1710000
WP ADN46464_17 1700001 1810000
WP ADN46464_18 1800001 1910000
WP ADN46464_19 1900001 2010000
WP ADN46464_20 2000001 2089378

Query Match 84.7%; Score 14.4; DB 12; Length 89378;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
|||||

Db 64028 GCCCAGCGTTGGCCGA 64043

RESULT 78

ADN47591_00/c

WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591

WP Fragment Name Begin End
WP ADN47591_00 1 110000
WP ADN47591_01 100001 210000
WP ADN47591_02 200001 310000
WP ADN47591_03 300001 410000
WP ADN47591_04 400001 510000
WP ADN47591_05 500001 610000
WP ADN47591_06 600001 710000
WP ADN47591_07 700001 810000
WP ADN47591_08 800001 910000
WP ADN47591_09 900001 1010000
WP ADN47591_10 1000001 1110000
WP ADN47591_11 1100001 1210000
WP ADN47591_12 1200001 1310000
WP ADN47591_13 1300001 1410000
WP ADN47591_14 1400001 1510000
WP ADN47591_15 1500001 1610000
WP ADN47591_16 1600001 1710000
WP ADN47591_17 1700001 1810000
WP ADN47591_18 1800001 1910000
WP ADN47591_19 1900001 2010000
WP ADN47591_20 2000001 2089378

ADN47591 standard; DNA; 2089378 BP.

XX AC ADN47591;

XX DT 01-JUL-2004 (first entry)

XX DE Thermococcus kodakaraensis KOD1 genome DNA sequence SeqID1469.

XX KW gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology;
KW ds; gene.

XX OS Thermococcus kodakaraensis.

XX PN WO2004022736-A1.

XX PD 18-MAR-2004.

XX PF 29-AUG-2003; 2003WO-IB003597.

XX PR 30-AUG-2002; 2002JP-00319011.

XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Imanaka T, Atomi H;

```
XX WPI; 2004-257583/24.
DR
XX
XX KW gene disruption; gene targeting; marker gene; transformation;
XX KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
XX KW gene structure; gene function; enzyme activity; medicine;
XX KW forensic science; food; drug inspection; molecular biology; immunology;
XX KW ds; gene.
XX OS Thermococcus kodakaraensis.
XX PN WO2004022736-A1.
XX PD 18-MAR-2004.
XX PF 29-AUG-2003; 2003WO-IB003597.
XX PR 30-AUG-2002; 2002JP-00319011.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Imanaka T, Atomi H;
XX DR WPI; 2004-257583/24.
XX PT Method for disrupting targeted gene in genome of organism particularly
XX PT thermotable bacterium and with genome chips for analysis, applicable in
XX PT studying gene structure and functions.
XX PS Claim 8; SEQ ID NO 1087; 598pp; Japanese.
XX SQ Sequence 2089378 BP; 499376A; 542308C; 544081G; 503594T; 0U; 190ther;

Query Match 84.7%; Score 14.4; DB 12; Length 110000;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGA 16
Db 25350 GCCCAGCTTGGCCGA 25335

RESULT 79
ADN47209 00/c
WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209
WP Fragment Name Begin End
WP ADN47209_00 1 110000
WP ADN47209_01 100001 210000
WP ADN47209_02 200001 310000
WP ADN47209_03 300001 410000
WP ADN47209_04 400001 510000
WP ADN47209_05 500001 610000
WP ADN47209_06 600001 710000
WP ADN47209_07 700001 810000
WP ADN47209_08 800001 910000
WP ADN47209_09 900001 1010000
WP ADN47209_10 1000001 1110000
WP ADN47209_11 1100001 1210000
WP ADN47209_12 1200001 1310000
WP ADN47209_13 1300001 1410000
WP ADN47209_14 1400001 1510000
WP ADN47209_15 1500001 1610000
WP ADN47209_16 1600001 1710000
WP ADN47209_17 1700001 1810000
WP ADN47209_18 1800001 1910000
WP ADN47209_19 1900001 2010000
WP ADN47209_20 2000001 2089378
ID ADN47209 standard; DNA; 2089378 BP.
XX
XX ADN47209;
XX AC
XX DT 01-JUL-2004 (first entry)
XX XX Thermococcus kodakaraensis KOD1 genome DNA sequence SeqID1087.
XX DE
XX XX
```

```
KW gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology;
KW ds; gene.
XX OS Thermococcus kodakaraensis.
XX PN WO2004022736-A1.
XX PD 18-MAR-2004.
XX PF 29-AUG-2003; 2003WO-IB003597.
XX PR 30-AUG-2002; 2002JP-00319011.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Imanaka T, Atomi H;
XX DR WPI; 2004-257583/24.
XX PT Method for disrupting targeted gene in genome of organism particularly
XX PT thermotable bacterium and with genome chips for analysis, applicable in
XX PT studying gene structure and functions.
XX PS Claim 8; SEQ ID NO 1087; 598pp; Japanese.
XX SQ Sequence 2089378 BP; 499376A; 542308C; 544081G; 503594T; 0U; 190ther;

Query Match 84.7%; Score 14.4; DB 12; Length 110000;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGA 16
Db 25350 GCCCAGCTTGGCCGA 25335

RESULT 80
ADN47960 00/c
WP Sequence split into 21 fragments LOCUS ADN47960 Accession Adn47960
WP Fragment Name Begin End
WP ADN47960_00 1 110000
WP ADN47960_01 100001 210000
WP ADN47960_02 200001 310000
WP ADN47960_03 300001 410000
WP ADN47960_04 400001 510000
WP ADN47960_05 500001 610000
WP ADN47960_06 600001 710000
WP ADN47960_07 700001 810000
WP ADN47960_08 800001 910000
WP ADN47960_09 900001 1010000
```

```
WP ADN47960_10 1000001 1110000
WP ADN47960_11 1100001 1210000
WP ADN47960_12 1200001 1310000
WP ADN47960_13 1300001 1410000
WP ADN47960_14 1400001 1510000
WP ADN47960_15 1500001 1610000
WP ADN47960_16 1600001 1710000
WP ADN47960_17 1700001 1810000
WP ADN47960_18 1800001 1910000
WP ADN47960_19 1900001 2010000
WP ADN47960_20 2000001 2089378
ID ADN47960 standard; DNA; 2089378 BP.
XX AC
XX ADN47960;
XX
XX 01-JUL-2004 (first entry)
XX
XX Thermococcus kodakaraensis KOD1 genome DNA sequence SeqID1838.
XX
XX gene disruption; gene targeting; marker gene; transformation;
XX homologous recombination; hyperthermostable archaeobacterium; KOD1;
XX gene structure; gene function; enzyme activity; medicine;
XX forensic science; food; drug inspection; molecular biology; immunology;
XX ds; gene.
XX
XX Thermococcus kodakaraensis.
XX
XX OS
XX WO2004022736-A1.
XX
XX 18-MAR-2004.
XX
XX 29-AUG-2003; 2003WO-IB003597.
XX
XX 30-AUG-2002; 2002JP-00319011.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Inanaka T, Atomí H;
XX
XX WPI; 2004-257583/24.
XX
XX Method for disrupting targeted gene in genome of organism particularly
XX thermotable bacterium and with genome chips for analysis, applicable in
XX studying gene structure and functions.
XX
XX Example 1; SEQ ID NO 1838; 598pp; Japanese.
XX
XX This invention relates to a novel method for targeting disruption of an
XX arbitrary gene in a genome of an organism which comprises providing the
XX whole sequential data of the genome of such organism, selecting at least
XX 1 arbitrary region in the sequence, providing a vector that contains a
XX sequence homologous with the selected region and a marker gene,
XX transformation, and homologous recombination. The genome is preferably
XX the genome of a hyperthermostable archaeobacterium, particularly
XX Thermococcus kodakaraensis KOD1. The method is for targeting the
XX disruption of a gene in the genome of an organism, which is applicable in
XX studying gene structure and functions as well as enzyme activities of
XX encoded proteins and useful in medicine, forensic science, food or drug
XX inspection, molecular biology and immunology. With this method, the
XX disruption of a gene at an arbitrary position in a genome can be achieved
XX efficiently and reliably. The present position is that of the genomic DNA
XX sequence of Thermococcus kodakaraensis (KOD1) which was derived during
XX use of the method of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2089378 BP; 499376A; 542308C; 544081G; 503594T; 0U; 190Other;
XX
XX Query Match 84.7%; Score 14.4; DB 12; Length 110000;
XX Best Local Similarity 93.8%; Pred. No. 1.5e+03;
XX Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 GCCACGCTTGGCGGA 16
Db 25350 GCCACGCTTGGCGGA 25335

RESULT 81
ABD32843
ID ABD32843 standard; DNA; 326002 BP.
XX AC
XX ABD32843;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human cancer-associated genomic DNA HD17-014.
XX
XX Human; ds; cancer-associated protein; gene; cytostatic; cancer;
XX leukaemia; lymphoma; CAP.
XX
XX Homo sapiens.
XX
XX WO2004074320-A2.
XX
XX 02-SEP-2004.
XX
XX 17-FEB-2004; 2004WO-US004730.
XX
XX 14-FEB-2003; 2003US-00367094.
XX
XX 14-MAR-2003; 2003US-00388838.
XX
XX 15-APR-2003; 2003US-00417375.
XX
XX 13-JUN-2003; 2003US-00461862.
XX
XX 15-SEP-2003; 2003US-00663431.
XX
XX 15-DEC-2003; 2003US-00737318.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Morris DW, Malandro MS;
XX
XX WPI; 2004-652914/63.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful
XX for diagnosing, preventing or treating cancers, especially lymphoma and
XX leukemia, or in screening for agents that modulate cancer.
XX
XX claim 16; seqid 506; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
XX contiguous nucleotides of any of the 233 polynucleotide sequences given
XX in the specification, or its complement. The nucleic acids encode cancer-
XX associated proteins. Also included are an expression vector comprising
XX the isolated nucleic acid cited above, a host cell comprising the above
XX recombinant nucleic acid or expression vector, a microarray for detecting
XX a cancer-associated (CA) nucleic acid comprising at least one probe
XX comprising at least 10 contiguous nucleotides of any of the above-
XX mentioned nucleotide sequences, an isolated polypeptide (encoded within
XX an open reading frame of a CA sequence selected from any of the 95
XX polynucleotide sequences as mentioned in the specification, or its
XX complement), an isolated antibody, (or its antigen binding fragment) that
XX binds to the above polypeptide, a hybridoma that produces the above
XX monoclonal antibody, a pharmaceutical composition comprising the above
XX antibody and a pharmaceutical excipient, a kit for detecting cancer
XX cells (comprising the antibody cited above, methods for diagnosing cancer
XX or for detecting the presence or absence of cancer cells in an
XX individual, a method for inhibiting a growth of cancer cells in an
XX individual, a method for delivering a therapeutic agent to cancer cells
XX in an individual, an electronic library comprising the above
XX polynucleotide or polypeptide (or their fragments), methods of screening
XX for anticancer activity or for a bioactive agent capable of modulating
XX the activity of a CA protein (CAP), methods for detecting cancer
XX associated with expression of a polypeptide in a test cell sample, a
XX method for treating cancers and a method for inhibiting the expression of
XX CA gene in a cell. The composition and methods are useful for detecting,
XX diagnosing, preventing and treating cancers, especially lymphoma and
XX leukaemia. These may also be used in screening for agents that modulate
```

CC cancer. The present sequence is a human CAP genomic sequence. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 326002 BP; 84319 A; 72970 C; 76829 G; 91884 T; 0 U; 0 Other;
 Query Match 84.7%; Score 14.4; DB 13; Length 326002;
 Best Local Similarity 93.8%; Pred. No. 1.5e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
 ||||| ||||| |||||
 Db 110753 CCCACCTTGCCGAG 110768

RESULT 82
 ABN23393/c
 ID ABN23393 standard; cDNA; 338 BP.
 XX AC ABN23393;
 XX DT 24-JUN-2002 (first entry)
 XX DE Human ORFX polynucleotide sequence SEQ ID NO:15263.
 XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200192523-A2.
 XX PD 06-DEC-2001.
 XX PF 29-MAY-2001; 2001WO-US010836.
 XX PR 30-MAY-2000; 2000US-0206132P.
 XX PR 29-AUG-2000; 2000US-0228716P.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA, Leach MD;
 XX DR WPI; 2002-106308/14.
 XX P-PSDB; ABP07641.
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 XX preventing and treating cardiovascular disease, neurodegenerative,
 XX hyperproliferative disorders and autoimmune disorders.
 XX PS Disclosure; SEQ ID NO 15263; 1037pp; English.
 XX

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP0010 to ABP1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 338 BP; 66 A; 98 C; 108 G; 64 T; 0 U; 2 Other;
 Query Match 82.4%; Score 14; DB 6; Length 338;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCC 14
 ||||| ||||| |||||
 Db 196 GCCACGCTTGCC 183

RESULT 83
 ACH41474/c
 ID ACH41474 standard; cDNA; 458 BP.
 XX AC ACH41474;
 XX DT 13-OCT-2003 (first entry)
 XX DE Human foetal brain cDNA #2841.
 XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.
 XX OS Homo sapiens.
 XX PN US2003073623-A1.
 XX PD 17-APR-2003.
 XX PF 30-JUL-2001; 2001US-00918995.
 XX PR 30-JUL-2001; 2001US-00918995.
 XX PA (DRMA/) DRMANAC R T.
 PA (LABA/) LABAT I.
 PA (STAC/) STACHE-CRAIN B.
 PA (DICK/) DICKSON M C.
 PA (JONE/) JONES L W.
 XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 XX WPI; 2003-615964/58.
 XX New polynucleotide sequences obtained from various cDNA libraries, useful
 XX as hybridization probes, as oligomers for PCR, for chromosome and gene
 XX mapping, in the recombinant production of protein, or in generating
 XX antisense DNA or RNA.
 XX PS Claim 1; SEQ ID NO 28686; 44pp; English.
 XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of

CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The present sequence
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX
SQ Sequence 458 BP; 101 A; 104 C; 124 G; 125 T; 0 U; 4 Other;
Query Match 82.4%; Score 14; DB 9; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTGGCCGA 16
| | | | | | | | | | | | | | | |
Db 358 CCACGCTGGCCGA 345

RESULT 84
ADN12918/c
ID ADN12918 standard; cDNA; 508 BP.
XX
AC ADN12918;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human prostate/colon/lung/breast cancer-related cDNA 433, SEQ:433.
XX
KW Human; cancer; tumour; prostate cancer; colon cancer; lung cancer;
KW breast cancer; drug screening; diagnosis; prognosis; prevention;
KW gene mapping; tissue typing; tissue profiling; cytostatic; gene therapy;
KW ss.
XX
OS Homo sapiens.
XX
PN WO2004039943-A2.
XX
PD 13-MAY-2004.
XX
PF 16-MAY-2003; 2003WO-US015465.
XX
PR 17-MAY-2002; 2002US-0381533P.
PR 04-FEB-2003; 2003US-0445222P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Scott EM, Lamson G, Kaseam A, Zhang G, Sakamoto D, Garcia PD;
XX
DR WPI; 2004-376173/35.
XX

XX
PT New isolated polynucleotides, useful for gene mapping or tissue typing or
PT profiling, as diagnostic reagents, and for preventing or treating cancer,
PT e.g. prostate, colon, or breast cancer.
XX
PS Claim 2; SEQ ID NO 433; 190pp; English.

XX
CC The invention relates to nucleic acids (ADN12486-ADN13970) isolated from
CC human prostate, colon, lung and breast cancer cDNA libraries, and to 57
CC proteins (ADN13971-ADN14027) encoded by a subset of these cDNA sequences
CC (ADN13914-ADN13970). The invention also relates to vectors and host cells
CC comprising a nucleic acid of the invention; a method for the recombinant
CC production of a protein of the invention; an antibody specific for a
CC protein of the invention; a polynucleotide library comprising at least
CC one nucleic acid sequence of the invention; a method for detecting a
CC cancerous cell by PCR or probe hybridization; inhibiting a cancerous
CC phenotype (particularly aberrant proliferation) of a cell; a method of
CC identifying an agent that modulates the biological activity of a gene
CC product differentially expressed in a cancerous cell compared with a
CC normal cell; and a method of treating a cancer patient using the agent
CC identified. The nucleic acids and polypeptides can be used to diagnose,
CC prognosis, treat or prevent cancers such as prostate, colon, lung or
CC breast cancer, and can also be used to screen for drugs for the treatment
CC of cancer. The nucleic acids can also be used for gene mapping, tissue

CC typing and tissue profiling. The present sequence represents a
CC specifically claimed cancer-related cDNA of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 508 BP; 115 A; 127 C; 156 G; 110 T; 0 U; 0 Other;

Query Match 82.4%; Score 14; DB 12; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTGGCCGA 16
| | | | | | | | | | | | | | | |
Db 238 CCACGCTGGCCGA 225

RESULT 85
ADN13857/c
ID ADN13857 standard; cDNA; 706 BP.
XX
AC ADN13857;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human prostate/colon/lung/breast cancer-related cDNA 1372, SEQ:1372.
XX
KW Human; cancer; tumour; prostate cancer; colon cancer; lung cancer;
KW breast cancer; drug screening; diagnosis; prognosis; prevention;
KW gene mapping; tissue typing; tissue profiling; cytostatic; gene therapy;
KW ss.
XX
OS Homo sapiens.
XX
PN WO2004039943-A2.
XX
PD 13-MAY-2004.
XX
PF 16-MAY-2003; 2003WO-US015465.
XX
PR 17-MAY-2002; 2002US-0381533P.
PR 04-FEB-2003; 2003US-0445222P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Scott EM, Lamson G, Kaseam A, Zhang G, Sakamoto D, Garcia PD;
XX
DR WPI; 2004-376173/35.
XX

XX
PT New isolated polynucleotides, useful for gene mapping or tissue typing or
PT profiling, as diagnostic reagents, and for preventing or treating cancer,
PT e.g. prostate, colon, or breast cancer.
XX
PS Claim 2; SEQ ID NO 1372; 190pp; English.

XX
CC The invention relates to nucleic acids (ADN12486-ADN13970) isolated from
CC human prostate, colon, lung and breast cancer cDNA libraries, and to 57
CC proteins (ADN13971-ADN14027) encoded by a subset of these cDNA sequences
CC (ADN13914-ADN13970). The invention also relates to vectors and host cells
CC comprising a nucleic acid of the invention; a method for the recombinant
CC production of a protein of the invention; an antibody specific for a
CC protein of the invention; a polynucleotide library comprising at least
CC one nucleic acid sequence of the invention; a method for detecting a
CC cancerous cell by PCR or probe hybridization; inhibiting a cancerous
CC phenotype (particularly aberrant proliferation) of a cell; a method of
CC identifying an agent that modulates the biological activity of a gene
CC product differentially expressed in a cancerous cell compared with a
CC normal cell; and a method of treating a cancer patient using the agent
CC identified. The nucleic acids and polypeptides can be used to diagnose,
CC prognosis, treat or prevent cancers such as prostate, colon, lung or
CC breast cancer, and can also be used to screen for drugs for the treatment
CC of cancer. The nucleic acids can also be used for gene mapping, tissue
CC typing and tissue profiling. The present sequence represents a

CC specifically claimed cancer-related cDNA of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 706 BP; 170 A; 176 C; 206 G; 150 T; 0 U; 4 Other;
Query Match 82.4%; Score 14; DB 12; Length 706;
Best Local Similarity 100.0%; Pred. No. 2.2e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CCACGCTTGGCCGA 16
Db 238 CCACGCTTGGCCGA 225

RESULT 86

AAC47788

ID AAC47788 standard; DNA; 1234 BP.

XX AAC47788;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55122.

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0147955P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 30-AUG-1999; 99US-0151438P.
PR 31-AUG-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 07-SEP-1999; 99US-0153070P.
PR 10-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 15-SEP-1999; 99US-0154039P.
PR 16-SEP-1999; 99US-0154779P.
PR 20-SEP-1999; 99US-0155139P.
PR 22-SEP-1999; 99US-0155486P.
PR 23-SEP-1999; 99US-0155659P.
PR 24-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 29-SEP-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157753P.
PR 05-OCT-1999; 99US-0157865P.
PR 06-OCT-1999; 99US-0158039P.
PR 07-OCT-1999; 99US-0158232P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160769P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.4%; Score 14; DB 3; Length 1234;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCC 14
|||||
Db 563 GCCCAGCTTGGCC 576

RESULT 87

ADN46845 16/c
Continuation (17 of 21) of ADN46845 from base 1600001 (Thermococcus kodakaraensis KOD1 Gen
WP Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845
WP Fragment Name Begin End
WP ADN46845_00 1 110000
WP ADN46845_01 100001 210000
WP ADN46845_02 200001 310000
WP ADN46845_03 300001 410000
WP ADN46845_04 400001 510000
WP ADN46845_05 500001 610000
WP ADN46845_06 600001 710000
WP ADN46845_07 700001 810000
WP ADN46845_08 800001 910000
WP ADN46845_09 900001 1010000
WP ADN46845_10 1000001 1110000
WP ADN46845_11 1100001 1210000
WP ADN46845_12 1200001 1310000
WP ADN46845_13 1300001 1410000
WP ADN46845_14 1400001 1510000
WP ADN46845_15 1500001 1610000
WP ADN46845_16 1600001 1710000
WP ADN46845_17 1700001 1810000
WP ADN46845_18 1800001 1910000
WP ADN46845_19 1900001 2010000
WP ADN46845_20 2000001 2089378

Query Match 82.4%; Score 14; DB 12; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGA 16
|||||
Db 53895 CCACGCTTGGCCGA 53882

RESULT 88

ADN47591_04
Continuation (5 of 21) of ADN47591 from base 400001 (Thermococcus kodakaraensis KOD1 Gen
WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591
WP Fragment Name Begin End
WP ADN47591_00 1 110000
WP ADN47591_01 100001 210000
WP ADN47591_02 200001 310000
WP ADN47591_03 300001 410000
WP ADN47591_04 400001 510000
WP ADN47591_05 500001 610000
WP ADN47591_06 600001 710000
WP ADN47591_07 700001 810000
WP ADN47591_08 800001 910000
WP ADN47591_09 900001 1010000
WP ADN47591_10 1000001 1110000
WP ADN47591_11 1100001 1210000
WP ADN47591_12 1200001 1310000
WP ADN47591_13 1300001 1410000
WP ADN47591_14 1400001 1510000
WP ADN47591_15 1500001 1610000
WP ADN47591_16 1600001 1710000
WP ADN47591_17 1700001 1810000
WP ADN47591_18 1800001 1910000
WP ADN47591_19 1900001 2010000
WP ADN47591_20 2000001 2089378

Query Match 82.4%; Score 14; DB 12; Length 110000;

Best Local Similarity 100.0%; Pred. No. 2.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGA 16
| | | | | | | | | |
DB 35483 CCACGCTTGGCCGA 35496

RESULT 89

ADN46123_16/c
Continuation (17 of 21) of ADN46123 from base 1600001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123

WP	Fragment Name	Begin	End
WP	ADN46123_00	1	110000
WP	ADN46123_01	100001	210000
WP	ADN46123_02	200001	310000
WP	ADN46123_03	300001	410000
WP	ADN46123_04	400001	510000
WP	ADN46123_05	500001	610000
WP	ADN46123_06	600001	710000
WP	ADN46123_07	700001	810000
WP	ADN46123_08	800001	910000
WP	ADN46123_09	900001	1010000
WP	ADN46123_10	1000001	1110000
WP	ADN46123_11	1100001	1210000
WP	ADN46123_12	1200001	1310000
WP	ADN46123_13	1300001	1410000
WP	ADN46123_14	1400001	1510000
WP	ADN46123_15	1500001	1610000
WP	ADN46123_16	1600001	1710000
WP	ADN46123_17	1700001	1810000
WP	ADN46123_18	1800001	1910000
WP	ADN46123_19	1900001	2010000
WP	ADN46123_20	2000001	2089378

Query Match 82.4%; Score 14; DB 12; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGA 16
| | | | | | | | | |
DB 53895 CCACGCTTGGCCGA 53882

RESULT 90

ADN47209_04
Continuation (5 of 21) of ADN47209 from base 400001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209

WP	Fragment Name	Begin	End
WP	ADN47209_00	1	110000
WP	ADN47209_01	100001	210000
WP	ADN47209_02	200001	310000
WP	ADN47209_03	300001	410000
WP	ADN47209_04	400001	510000
WP	ADN47209_05	500001	610000
WP	ADN47209_06	600001	710000
WP	ADN47209_07	700001	810000
WP	ADN47209_08	800001	910000
WP	ADN47209_09	900001	1010000
WP	ADN47209_10	1000001	1110000
WP	ADN47209_11	1100001	1210000
WP	ADN47209_12	1200001	1310000
WP	ADN47209_13	1300001	1410000
WP	ADN47209_14	1400001	1510000
WP	ADN47209_15	1500001	1610000
WP	ADN47209_16	1600001	1710000
WP	ADN47209_17	1700001	1810000
WP	ADN47209_18	1800001	1910000
WP	ADN47209_19	1900001	2010000
WP	ADN47209_20	2000001	2089378

Query Match 82.4%; Score 14; DB 12; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGA 16
| | | | | | | | | |
DB 35483 CCACGCTTGGCCGA 35496

RESULT 91

ADN46464_16/c
Continuation (17 of 21) of ADN46464 from base 1600001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN46464 Accession Adn46464

WP	Fragment Name	Begin	End
WP	ADN46464_00	1	110000
WP	ADN46464_01	100001	210000
WP	ADN46464_02	200001	310000
WP	ADN46464_03	300001	410000
WP	ADN46464_04	400001	510000
WP	ADN46464_05	500001	610000
WP	ADN46464_06	600001	710000
WP	ADN46464_07	700001	810000
WP	ADN46464_08	800001	910000
WP	ADN46464_09	900001	1010000
WP	ADN46464_10	1000001	1110000
WP	ADN46464_11	1100001	1210000
WP	ADN46464_12	1200001	1310000
WP	ADN46464_13	1300001	1410000
WP	ADN46464_14	1400001	1510000
WP	ADN46464_15	1500001	1610000
WP	ADN46464_16	1600001	1710000
WP	ADN46464_17	1700001	1810000
WP	ADN46464_18	1800001	1910000
WP	ADN46464_19	1900001	2010000
WP	ADN46464_20	2000001	2089378

Query Match 82.4%; Score 14; DB 12; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGA 16
| | | | | | | | | |
DB 53895 CCACGCTTGGCCGA 53882

RESULT 92

ADN47960_04
Continuation (5 of 21) of ADN47960 from base 400001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN47960 Accession Adn47960

WP	Fragment Name	Begin	End
WP	ADN47960_00	1	110000
WP	ADN47960_01	100001	210000
WP	ADN47960_02	200001	310000
WP	ADN47960_03	300001	410000
WP	ADN47960_04	400001	510000
WP	ADN47960_05	500001	610000
WP	ADN47960_06	600001	710000
WP	ADN47960_07	700001	810000
WP	ADN47960_08	800001	910000
WP	ADN47960_09	900001	1010000
WP	ADN47960_10	1000001	1110000
WP	ADN47960_11	1100001	1210000
WP	ADN47960_12	1200001	1310000
WP	ADN47960_13	1300001	1410000
WP	ADN47960_14	1400001	1510000
WP	ADN47960_15	1500001	1610000
WP	ADN47960_16	1600001	1710000
WP	ADN47960_17	1700001	1810000
WP	ADN47960_18	1800001	1910000
WP	ADN47960_19	1900001	2010000
WP	ADN47960_20	2000001	2089378

Query Match 82.4%; Score 14; DB 12; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      3 CCACGCTGGCCGA 16
Db      35483 CCACGCTGGCCGA 35496

RESULT 93
ABQ81848/c
ID      ABQ81848 standard; DNA; 349980 BP.
XX
AC      ABQ81848;
XX
DT      19-NOV-2002 (first entry)
XX
DE      Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1104.
XX
KW      Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW      antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW      identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW      rotavirus; food composition; pharmaceutical composition; gene; ds.
XX
OS      Bifidobacterium longum.
OS      Synthetic.
XX
XX      EPI227152-A1.
XX
XX      31-JUL-2002.
XX
XX      30-JAN-2001; 2001EP-00102050.
XX
XX      30-JAN-2001; 2001EP-00102050.
XX
XX      (NEST ) SOC PROD NESTLE SA.
XX
XX      WPI; 2002-668397/72.
XX
PT      Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT      a probe or primer for detecting and/or identifying Bifidobacterium longum
PT      in a biological sample.
XX
XX      Disclosure; SEQ ID NO 1104; 80pp; English.
XX
CC      The present invention describes a polynucleotide (I) comprising a
CC      sequence of a Bifidobacterium genome selected from the nucleotide
CC      sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
CC      least 90% identity or which hybridises with the sequences given in
CC      ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
CC      fusion protein, comprising a sequence selected from 1097 sequences given
CC      in ABP65258 to ABP63354 ligated in frame to a polynucleotide encoding a
CC      heterologous polypeptide. (I) has antidiarrheic and antibacterial
CC      activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC      a probe) is useful for the detection and/or identification of
CC      Bifidobacterium longum in a biological sample. A carrier containing the
CC      lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM 1-2618) can be
CC      used for preventing and/or treating diarrhoea brought about by pathogenic
CC      bacteria and/or rotavirus. The carrier is a food composition selected
CC      from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC      products, ice-creams, fermented cereal based products, milk based
CC      powders, infant formula, pet food or a pharmaceutical composition
CC      selected from tablets, liquid bacterial suspensions, dried oral
CC      supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC      (I) is useful in DNA arrays or chips to carry out analysis of the
CC      expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
CC      Bifidobacterium related nucleotide sequences given in the Sequence
CC      Listing from the present invention but not mentioned further within the
CC      specification. N.B. The sequence data for this patent is not represented
CC      in the printed specification but is based on sequence information
CC      supplied by the European Patent Office
XX
SQ      Sequence 349980 BP; 69195 A; 106952 C; 106128 G; 67705 T; 0 U; 0 Other;
Query Match      82.4%; Score 14; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;
```

```
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      4 CACGCTTGGCCGAG 17
      |||||
Db      298538 CACGCTTGGCCGAG 298525

RESULT 94
ABQ81844
ID      ABQ81844 standard; DNA; 349980 BP.
XX
XX      ABQ81844;
AC      ABQ81844;
XX
DT      19-NOV-2002 (first entry)
XX
DE      Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1100.
XX
KW      Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW      antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW      identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW      rotavirus; food composition; pharmaceutical composition; gene; ds.
XX
OS      Bifidobacterium longum.
OS      Synthetic.
XX
XX      EPI227152-A1.
XX
XX      31-JUL-2002.
XX
XX      30-JAN-2001; 2001EP-00102050.
XX
XX      30-JAN-2001; 2001EP-00102050.
XX
XX      (NEST ) SOC PROD NESTLE SA.
XX
XX      WPI; 2002-668397/72.
XX
PT      Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT      a probe or primer for detecting and/or identifying Bifidobacterium longum
PT      in a biological sample.
XX
XX      Disclosure; SEQ ID NO 1100; 80pp; English.
XX
CC      The present invention describes a polynucleotide (I) comprising a
CC      sequence of a Bifidobacterium genome selected from the nucleotide
CC      sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
CC      least 90% identity or which hybridises with the sequences given in
CC      ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
CC      fusion protein, comprising a sequence selected from 1097 sequences given
CC      in ABP65258 to ABP63354 ligated in frame to a polynucleotide encoding a
CC      heterologous polypeptide. (I) has antidiarrheic and antibacterial
CC      activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC      a probe) is useful for the detection and/or identification of
CC      Bifidobacterium longum in a biological sample. A carrier containing the
CC      lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM 1-2618) can be
CC      used for preventing and/or treating diarrhoea brought about by pathogenic
CC      bacteria and/or rotavirus. The carrier is a food composition selected
CC      from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC      products, ice-creams, fermented cereal based products, milk based
CC      powders, infant formula, pet food or a pharmaceutical composition
CC      selected from tablets, liquid bacterial suspensions, dried oral
CC      supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC      (I) is useful in DNA arrays or chips to carry out analysis of the
CC      expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
CC      Bifidobacterium related nucleotide sequences given in the Sequence
CC      Listing from the present invention but not mentioned further within the
CC      specification. N.B. The sequence data for this patent is not represented
CC      in the printed specification but is based on sequence information
CC      supplied by the European Patent Office
XX
SQ      Sequence 349980 BP; 69975 A; 105045 C; 104394 G; 70566 T; 0 U; 0 Other;
Query Match      82.4%; Score 14; DB 6; Length 349980;
```

```

Best Local Similarity 100.0%; Pred. No. 2.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGGCTTGCC 14
Db 242514 GCCACGGCTTGCC 242527

RESULT 95
AAZ68079
ID AAZ68079 standard; DNA; 47 BP.
XX AC AAZ68079;
XX 10-SEP-2001 (first entry)
XX Human map-related biallelic marker SEQ ID NO:2426.
DE Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridization; identification; characterisation; diagnosis;
KW single nucleotide polymorphism; SNP; ds.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT variation replace(24,C)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
XX WO9954500-A2.
XX 28-OCT-1999.
XX
XX 21-APR-1999; 99WO-IB000822.
XX 21-APR-1998; 98US-0082614P.
XX 23-NOV-1998; 98US-0109732P.
XX (GEST ) GENSET.
XX
XX Cohen D, Blumenfeld M, Chumakov I;
XX WPI; 2000-013267/01.
XX
XX Novel biallelic markers used to construct a high density disequilibrium
XX map of the human genome.
XX Claim 3; Page 748; 2745pp; English.
XX
XX AAZ65654 to AAZ69578 represent human biallelic markers from the present
XX invention, which contain a polymorphic base at position 24 of their
XX nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
XX primers for the biallelic markers. The biallelic markers of the invention
XX have a variety of uses: they can be used for high density mapping of the
XX human genome, and in complex association studies and haplotyping studies
XX which are useful in determining the genetic basis for disease states.
XX Compositions and methods of the invention can also be useful for the
XX identification of the targets for the development of pharmaceutical
XX agents and diagnostic methods, as well as the characterisation of the
XX differential efficacious responses to and side effects from
XX pharmaceutical agents acting on a disease as well as other treatment.
XX N.B. The SEQ ID Nos 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
XX 3367, are not actually given a sequence in the Sequence Listing from the
XX present invention
XX
XX Sequence 47 BP; 8 A; 17 C; 11 G; 11 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 13.8; DB 3; Length 47;
XX Best Local Similarity 88.2%; Pred. No. 2.7e+03;
XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCACGGCTTGCCGAG 17
Db 27 GCCACGGCTTGCCGAG 43

RESULT 96
ABN78126
ID ABN78126 standard; cDNA; 228 BP.
XX AC ABN78126;
XX 08-JUL-2002 (first entry)
XX Human glycoprotein-like ORF3073 cDNA, SEQ ID NO:6145.
XX
XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
XX disease monitoring; cytokine; cell proliferation; cell differentiation;
XX immune modulation; haematopoiesis regulation; tissue growth;
XX angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; tumour inhibition; bodily characteristic; fertility;
XX behaviour; cancer; proliferative disorder; neurological disorder;
XX cardiovascular disease; immune system disorder; organ transplantation;
XX tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
XX hypothyroidism; cholesterol ester storage disease; infection; vulnery;
XX vasotropic; antipsoriatic; antidiabetic; cytoagulant; neurotropic;
XX neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
XX cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
XX dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200190366-A2.
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US017076.
XX 24-MAY-2000; 2000US-0206690P.
XX (CURA-) CURAGEN CORP.
XX
XX Leach MD, Shimkets RA;
XX WPI; 2002-106200/14.
XX P-PSDB; ABP34100.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and disorders related to organ
XX transplantation.
XX Claim 1; Page 1786; 2508pp; English.
XX
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX ABN79587 represent cDNAs encoding them. The invention also encompasses
XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX polynucleotides, the recombinant production of ORFX proteins, antibodies
XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
XX polypeptides, methods of screening for modulators of ORFX expression or
XX activity, and methods of screening individuals for a predisposition to an
XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX range of biological activities, such as cytokine, cell proliferation,
XX cell differentiation, immune modulation, haematopoiesis regulation,
XX tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
XX chemokinetic activity, haemostatic activity, tumour inhibition activity,
XX receptor/ligand, antiinflammatory activity, tumour inhibition activity,
XX and antiinfective activity, and may also be involved in the determination
XX of bodily characteristics, fertility and behaviour. ORFX proteins,
XX nucleic acids and antibodies may be used in the treatment of cancers,
XX other proliferative disorders such as psoriasis and benign tumours,
XX neurological disorders such as epilepsy and Alzheimer's disease,

```

CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ Sequence 228 BP; 42 A; 77 C; 80 G; 28 T; 0 U; 1 Other;
Query Match 81.2%; Score 13.8; DB 6; Length 228;
Best Local Similarity 88.2%; Pred. No. 2.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGGCCGAG 17
Db 65 GCGCAGCGCTTGGCCGAG 81
RESULT 97
AAD48755 AAD48755 standard; RNA; 247 BP.
XX AAD48755;
XX 07-MAR-2003 (first entry)
XX Human acyl CoA:diacylglycerol acetyltransferase (DGAT) RNA fragment #1.
XX Human; RNA binding molecule; acyl CoA:diacylglycerol acetyltransferase;
XX DGAT; ss.
XX Homo sapiens.
XX WO200281748-A1.
XX 17-OCT-2002.
XX 04-APR-2002; 2002WO-SE000677.
XX 05-APR-2001; 2001SE-00001218.
XX 05-APR-2001; 2001US-0281384P.
XX (BIOV-) BIOVITRUM AB.
XX Ekblom J;
XX WPI; 2003-058568/05.
XX Identifying RNA-binding molecule by predicting structure of RNA fragment,
XX synthesizing DNA fragment corresponding to predicted RNA structure,
XX performing reporter gene assay after placing the DNA upstream of reporter
XX gene.
XX Claim 12; Page 27; 35pp; English.
XX The present invention relates to a method of identifying RNA-binding
XX molecule comprising predicting the structure of RNA-fragment, selecting
XX suitable predicted RNA fragment with an individual stem, synthesising a
XX DNA-fragment corresponding to the RNA fragment, inserting the DNA
XX fragment in upstream proximity of reporter assay gene to form reporter
XX construct and performing a reporter gene assay which detects interaction
XX between a molecule to be tested for RNA-binding and RNA fragment of the
XX reporter construct. The method is useful for identifying an RNA binding
XX molecule. The present sequence is human acyl CoA:diacylglycerol-
XX acetyltransferase (DGAT) RNA fragment. This sequence is used to
XX illustrate the method of the invention

XX SQ Sequence 247 BP; 32 A; 74 C; 108 G; 0 T; 33 U; 0 Other;
Query Match 81.2%; Score 13.8; DB 10; Length 247;
Best Local Similarity 76.5%; Pred. No. 2.8e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGGCCGAG 17
Db 208 GCCCAGCGCUUGGCGCG 224
RESULT 98
ABX84111/c
ID ABX84111 standard; cDNA; 283 BP.
XX ABX84111;
XX 24-APR-2003 (first entry)
XX Corn ear-derived polynucleotide (cpd) #2571.
XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
XX structural gene; functional gene; regulatory gene;
XX corn ear-specific profile; gene transcription; gene expression;
XX hybrid plant; desirable trait expression; plant breeding program;
XX inheritance; desired characteristic; growth; development;
XX disease resistance; environmental adaptability; quality; yield;
XX multigene trait; plant; gene; ss.
XX Zea mays.
XX US6476212-B1.
XX 05-NOV-2002.
XX 14-MAY-1999; 99US-00313294.
XX 26-MAY-1998; 98US-0086722P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lalgudi RV, Ito LY, Sherman BK;
XX WPI; 2003-208840/20.
XX Novel purified corn-ear derived polynucleotide useful as hybridization
XX probe for detecting polynucleotide in sample, and for identifying,
XX evaluating, and altering desired characteristics associated with growth,
XX development.
XX Example; SEQ ID NO 2571; 390pp; English.
XX The present invention relates to the isolation of corn ear-derived
XX polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
XX and SATMON023. Some of the cdps uniquely identify structural, functional,
XX and regulatory genes of corn ear. The polynucleotide sequences are
XX useful for detecting cdps in a sample, for producing a corn ear-specific
XX profile of gene transcription, for detecting altered gene expression in
XX inbred or hybrid plants, and for screening several molecules for specific
XX binding to the polynucleotide. The cdps are useful to identify, isolate,
XX or extend identical or related corn-ear nucleic acid sequences from DNA
XX libraries, and in nucleic acid amplification or hybridisation techniques
XX to follow the expression of desirable traits through plant breeding
XX programs. Preferably, the cdps are used to identify, evaluate, alter, or
XX follow the inheritance of desired characteristics associated with growth
XX and development, disease resistance, environmental adaptability, quality,
XX and yield of corn. The cdps are also useful as molecular markers for
XX studying inheritance and multigene traits in a plant breeding program.
XX The cdps are useful for producing purified corn-ear polypeptides by
XX recombinant techniques. They are also useful in diagnostic assays to
XX detect or confirm conditions or diseases associated with abnormal levels
XX of cdp expression. ABX81541-ABX89140 represent corn ear-derived

CC polynucleotides (cpds) of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipidsEntry.html

XX SQ Sequence 283 BP; 82 A; 56 C; 73 G; 72 T; 0 U; 0 Other;
Query Match 81.2%; Score 13.8; DB 10; Length 283;
Best Local Similarity 88.2%; Pred. No. 2.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCAGCTTGCCGAG 17
||||| ||||| ||
Db 106 GCCCACCCTTGCCCAAG 90

RESULT 99
ADP95237/c
ID ADP95237 standard; cDNA; 290 BP.

XX AC ADP95237;

XX DT 09-SEP-2004 (first entry)

XX DE Cotton expressed sequence tag, EST, #4248.

XX KW Cotton; ss; EST: expressed sequence tag; plant; plant protection;
KW plant improvement; marker-assisted breeding.

XX OS Gossypium hirsutum; variety Nucleotides33B.

XX PN US2004123338-A1.

XX PD 24-JUN-2004.

XX PF 08-DEC-2000; 2000US-00732627.

XX PR 10-DEC-1999; 99US-0170255P.

XX PA (FINC/) FINCHER K L.

XX PI Fincher KL;

XX PS WPI; 2004-479807/45.

XX PT New substantially purified nucleic acid molecule that encodes a cotton
PT protein or its fragment, useful as molecular tool for the targeting and
PT isolation of novel genes for plant protection and improvement.

XX PS Claim 1; SEQ ID NO 4248; 30pp; English.

XX CC The invention relates to a substantially purified nucleic acid molecule
CC that encodes a cotton protein or its fragment comprising an EST
CC (expressed sequence tag) appearing as ADP9990-ADP95919. Also included
CC are a substantially purified cotton protein or its fragment encoded by a
CC nucleic acid molecule above and a transformed plant (having a nucleic
CC acid molecule which comprises: an exogenous promoter region which
CC functions in a plant cell to cause the production of a mRNA molecule; a
CC structural nucleic acid molecule comprising one of the ESTs or their
CC complements; a 3' non-translated sequence that functions in the plant
CC cell to cause termination of transcription and addition of polyadenylated
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC molecular tool for the targeting and isolation of novel genes for plant
CC protection and improvement. The ESTs are useful for developing new
CC strategies for understanding critical plant developmental and metabolic
CC pathways, for isolating genes and promoters, for identifying and mapping
CC the genes involved in developmental and metabolic pathways, and for
CC determining gene function. The cotton nucleic acid molecules are useful
CC as molecular tags to isolate genetic regions, isolate genes, map genes,
CC and determine gene function. The nucleic acid molecules are useful for
CC determining if genes are members of a particular gene family and for use
CC in marker-assisted breeding programs. The present sequence is one of the
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed

CC in the specification but are available in electronic format from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=20040123338.

XX SQ Sequence 290 BP; 81 A; 63 C; 67 G; 79 T; 0 U; 0 Other;

Query Match 81.2%; Score 13.8; DB 12; Length 290;
Best Local Similarity 88.2%; Pred. No. 2.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGAG 17
||||| ||||| |||||
Db 116 GCCCAGCTTGCCGAG 100

RESULT 100
ABX86649/c

ID ABX86649 standard; cDNA; 293 BP.

XX AC ABX86649;

XX DT 24-APR-2003 (first entry)

XX DE Corn ear-derived polynucleotide (cpd) #5109.

XX KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW structural gene; functional gene; regulatory gene;
KW corn ear-specific profile; gene transcription; gene expression;
KW hybrid plant; desirable trait expression; plant breeding program;
KW inheritance; desired characteristic; growth; development;
KW disease resistance; environmental adaptability; quality; yield;
KW multigene trait; plant; gene; ss.

XX OS Zea mays.

XX PN US6476212-B1.

XX PD 05-NOV-2002.

XX PF 14-MAY-1999; 99US-00313294.

XX PR 26-MAY-1998; 98US-0086722P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Lalgudi RV, Ito LY, Sherman BK;

XX DR WPI; 2003-208840/20.

XX PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with growth,
PT development.

XX PS Example; SEQ ID NO 5109; 390pp; English.

XX CC The present invention relates to the isolation of corn ear-derived
CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC and SATMON023. Some of the cdps uniquely identify structural, functional,
CC and regulatory genes of corn ear. The polynucleotide sequences are
CC useful for detecting cdps in a sample, for producing a corn ear-specific
CC profile of gene transcription, for detecting altered gene expression in
CC inbred or hybrid plants, and for screening several molecules for specific
CC binding to the polynucleotide. The cdps are useful to identify, isolate,
CC or extend identical or related corn-ear nucleic acid sequences from DNA
CC libraries, and in nucleic acid amplification or hybridisation techniques
CC to follow the expression of desirable traits through plant breeding
CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with growth
CC and development, disease resistance, environmental adaptability, quality,
CC and yield of corn. The cdps are also useful as molecular markers for
CC studying inheritance and multigene traits in a plant breeding program.
CC The cdps are useful for producing purified corn-ear polypeptides by
CC recombinant techniques. They are also useful in diagnostic assays to

CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cpds) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipdIDEntry.html

XX
 SQ Sequence 293 BP; 77 A; 62 C; 72 G; 75 T; 0 U; 7 Other;

Query Match 81.2%; Score 13.8; DB 10; Length 293;
 Best Local Similarity 88.2%; Pred. No. 2.8e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCCGAG 17
 |||||
 Db 180 GCCCAGCGCTTGGCCGAG 164

Search completed: June 10, 2006, 15:34:01
 Job time : 160.617 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:19:28 ; Search time 2097.8 Seconds
(without alignments)
453.155 Million cell updates/sec

Title: US-10-600-816-31
Perfect score: 17
Sequence: 1 gccacgctggccgag 17

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est3: *
3: gb_est4: *
4: gb_est5: *
5: gb_est6: *
6: gb_est2: *
7: gb_est2: *
8: gb_est7: *
9: gb_est8: *
10: gb_est9: *
11: gb_gss1: *
12: gb_gss2: *
13: gb_gss3: *
14: gb_gss4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	100.0	158	2	B1063311 IL3-UT011
2	17	100.0	307	7	B8818833 CM2-BN030
3	17	100.0	341	7	B8818817 CM2-BN030
4	17	100.0	367	1	A1112374 zn8elo.r
5	17	100.0	384	7	B8818835 CM2-BN030
6	17	100.0	396	7	B8818828 CM2-BN030
7	17	100.0	404	7	B8818840 CM2-BN030
8	17	100.0	424	7	B8818805 CM2-BN030
9	17	100.0	453	5	CB672930 fg8a05.y
10	17	100.0	458	7	B8818871 CM2-BN030
11	17	100.0	466	7	B8818859 CM2-BN030
12	17	100.0	467	7	B8818852 CM2-BN030
13	17	100.0	528	7	AW390096 CM3-ST017
14	17	100.0	571	3	B2553378 BP253378
15	17	100.0	577	4	CB152636 K-EST0209
16	17	100.0	578	3	BP267556 BP267556
17	17	100.0	580	3	BP265234 BP265234
18	17	100.0	580	3	BP267627 BP267627
19	17	100.0	580	3	BP268151 BP268151

20	17	100.0	582	3	BP256144
21	17	100.0	582	3	BP267918
22	17	100.0	582	3	BP363758
23	17	100.0	583	7	BE006062 RC2-BN012
24	17	100.0	584	3	BP256147
25	17	100.0	621	4	CB130950 K-EST0180
26	17	100.0	630	7	BE747107
27	17	100.0	637	3	BM843051 K-EST0120
28	17	100.0	699	7	BF568108
29	17	100.0	702	2	BE251131
30	17	100.0	851	3	BQ887352
31	17	100.0	865	4	CA454746
32	17	100.0	877	2	BN011054
33	17	100.0	899	3	BQ58341
34	17	100.0	901	2	BG831564
35	17	100.0	916	5	CF780868
36	17	100.0	918	4	BX372160
37	17	100.0	923	3	BUS43952
38	17	100.0	962	4	BX370558
39	17	100.0	1023	3	BQ067433
40	17	100.0	1125	2	BI193620
41	17	100.0	1830	6	CR597125
42	15.4	90.6	322	1	AA041289
43	15.4	90.6	363	4	BY153767
44	15.4	90.6	365	4	BY774337
45	15.4	90.6	365	7	BE818831
46	15.4	90.6	367	4	BY674396
47	15.4	90.6	383	4	BY674625
48	15.4	90.6	410	7	AW206921
49	15.4	90.6	449	7	BE818807
50	15.4	90.6	451	14	DU607602
51	15.4	90.6	494	3	BQ65719
52	15.4	90.6	566	9	DA945234
53	15.4	90.6	570	9	DA932540
54	15.4	90.6	577	5	CK783403
55	15.4	90.6	582	11	AZ259233
56	15.4	90.6	586	12	CE755799
57	15.4	90.6	593	9	DA940956
58	15.4	90.6	600	5	CF751855
59	15.4	90.6	601	8	CO196877
60	15.4	90.6	609	13	CW480200
61	15.4	90.6	669	2	BG293309
62	15.4	90.6	671	13	CW480201
63	15.4	90.6	689	9	DN265926
64	15.4	90.6	694	9	DN441215
65	15.4	90.6	721	8	CO196932
66	15.4	90.6	744	12	CC111772
67	15.4	90.6	749	13	CW576188
68	15.4	90.6	788	8	CR374717
69	15.4	90.6	797	7	BF700990
70	15.4	90.6	798	8	CO196873
71	15.4	90.6	802	11	BZ455483
72	15.4	90.6	816	13	CW616903
73	15.4	90.6	830	13	CW646415
74	15.4	90.6	851	4	CA495187
75	15.4	90.6	859	12	CC112512
76	15.4	90.6	914	12	CC851671
77	15.4	90.6	929	14	CNS02J9X
78	15.4	90.6	940	3	BQ685729
79	15.4	90.6	1058	10	DM341470
80	15.4	90.6	1117	9	DR125588
81	15.4	90.6	1156	7	BE729821
82	15.4	90.6	1263	5	CD256036
83	15.4	90.6	1388	6	CNS0E2L3
84	15.4	90.6	1454	6	CNS0GOND
85	15.4	90.6	1492	6	CNS0FYVC
86	15.4	90.6	1505	6	CNS0GM2S
87	15.4	90.6	1512	6	CNS0C3DQ
88	15.4	90.6	1518	6	CNS0EXOO
89	15.4	90.6	1962	14	DQ049847
90	15.4	90.6	4208	6	BC068256
91	15	88.2	198	2	BI529325
92	15	88.2	278	1	AI613865

BP256144	BP256144
BP267918	BP267918
BP363758	BP363758
BE006062	RC2-BN012
BP256147	BP256147
CB130950	K-EST0180
BE747107	601580680
BM843051	K-EST0120
BF568108	602183908
BE251131	602364932
BQ887352	AGENCOURT
CA454746	AGENCOURT
BN011054	603634744
BQ58341	AGENCOURT
BG831564	602765819
CF780868	AGENCOURT
BX372160	BX372160
BUS43952	AGENCOURT
BX370558	BX370558
BQ067433	AGENCOURT
BI193620	602946519
CR597125	full-leng
AA041289	zf08d003.8
BY153767	BY153767
BY774337	BY774337
BE818831	CM2-BN030
BY674396	BY674396
BY674625	BY674625
AW206921	UI-H-B11-
BE818807	CM2-BN030
DU607602	OO_Ba010
BQ65719	QHB23M12
DA945234	DA945234
DA932540	DA932540
CK783403	UI-D-GC1-
AZ259233	RPCI-23-1
CE755799	tigr-gss-
DA940956	DA940956
CF751855	UI-D-GC0-
CO196877	GEOL_2_F0
CW480200	fsbb001f2
BG293309	HNC60-1-H
CW480201	fsbb001f2
DN265926	LIB30320
DN441215	LIB5338-1
CO196932	GEOL_2_S0
CC111772	NDL_46M15
CW576188	OA_ABA010
CR374717	CR374717
BF700990	602127938
CO196873	GEOL_2_E0
BZ455483	BONIS66TR
CW616903	OA_ABA016
CW646415	OA_ABA018
CA495187	AGENCOURT
CC112512	NDL_37C12
CC851671	NDL_55M12
AUI9950	Tetraodon
BQ685729	AGENCOURT
DM341470	PP_LEC000
DR125588	49128899
BE729821	601564735
CD256036	AGENCOURT
CNS0E2L3	Tetraodon
CNS0GOND	Tetraodon
CNS0FYVC	Tetraodon
CNS0GM2S	Tetraodon
CNS0C3DQ	Tetraodon
CNS0EXOO	Tetraodon
DQ049847	Homo sapi
BC068256	Homo sapi
BI529325	1024097G0
AI613865	vg88d01.y

93	15	88.2	309	10	DR992377	DR992377 Mdlr7011A	c 166	14.4	84.7	341	7	AW160798
94	15	88.2	322	7	AV850077	AV850077 AV850077	c 167	14.4	84.7	353	5	CI305825
95	15	88.2	368	4	CA661056	CA661056 w1m1.pk00	c 168	14.4	84.7	355	2	BI148829
96	15	88.2	419	3	BW421246	BW421246	c 169	14.4	84.7	356	2	BI038890
97	15	88.2	438	8	CN899452	CN899452 010706ABA	c 170	14.4	84.7	364	4	C90338
98	15	88.2	484	3	BW366787	BW366787 BW366787	c 171	14.4	84.7	365	5	CK560001
99	15	88.2	492	3	BW178352	BW178352 BW178352	c 172	14.4	84.7	368	1	AA775070
100	15	88.2	517	3	BW203028	BW203028 BW203028	c 173	14.4	84.7	370	1	AI688285
101	15	88.2	518	7	AV884732	AV884732 AV884732	c 174	14.4	84.7	373	3	BQ738777
102	15	88.2	521	4	CA689031	CA689031 w1m96.pk0	c 175	14.4	84.7	379	1	AA994272
103	15	88.2	524	3	BW426366	BW426366 BW426366	c 176	14.4	84.7	384	8	CV127754
104	15	88.2	526	3	BW382890	BW382890 BW382890	c 177	14.4	84.7	392	2	BI534558
105	15	88.2	528	3	BW308736	BW308736 BW308736	c 178	14.4	84.7	392	8	CV095612
106	15	88.2	528	3	BW309783	BW309783 BW309783	c 179	14.4	84.7	404	3	BM839442
107	15	88.2	528	3	BW405257	BW405257 BW405257	c 180	14.4	84.7	409	1	AA159571
108	15	88.2	529	3	BW188031	BW188031 BW188031	c 181	14.4	84.7	414	5	CJ061166
109	15	88.2	534	8	CO903318	CO903318 Mddb5022e	c 182	14.4	84.7	429	7	BB875898
110	15	88.2	536	3	BW389352	BW389352 BW389352	c 183	14.4	84.7	431	8	CO872640
111	15	88.2	542	3	BW120319	BW120319 BW120319	c 184	14.4	84.7	445	10	N38821
112	15	88.2	543	3	BP024363	BP024363 BP024363	c 185	14.4	84.7	461	9	CV766672
113	15	88.2	554	3	BW472578	BW472578 BW472578	c 186	14.4	84.7	463	1	AA130015
114	15	88.2	554	7	AV854521	AV854521 AV854521	c 187	14.4	84.7	471	5	CI149721
115	15	88.2	554	7	AV976003	AV976003 AV976003	c 188	14.4	84.7	471	14	DU819050
116	15	88.2	557	3	BP019833	BP019833 BP019833	c 189	14.4	84.7	473	9	CV766470
117	15	88.2	560	3	BP019970	BP019970 BP019970	c 190	14.4	84.7	473	9	CV766497
118	15	88.2	561	3	BW306390	BW306390 BW306390	c 191	14.4	84.7	473	9	CV766740
119	15	88.2	565	3	BW095576	BW095576 BW095576	c 192	14.4	84.7	482	8	CV294645
120	15	88.2	565	7	AV962135	AV962135 AV962135	c 193	14.4	84.7	492	7	BV758506
121	15	88.2	566	7	AV864889	AV864889 AV864889	c 194	14.4	84.7	508	4	BX507570
122	15	88.2	567	3	BW315245	BW315245 BW315245	c 195	14.4	84.7	526	1	BA039719
123	15	88.2	571	3	BP015522	BP015522 BP015522	c 196	14.4	84.7	526	7	BF267413
124	15	88.2	572	3	BW411756	BW411756 BW411756	c 197	14.4	84.7	530	8	CO650574
125	15	88.2	573	3	BW121837	BW121837 BW121837	c 198	14.4	84.7	533	12	CC734250
126	15	88.2	573	7	AV996173	AV996173 AV996173	c 199	14.4	84.7	533	13	CZ366302
127	15	88.2	575	3	BW448598	BW448598 BW448598	c 200	14.4	84.7	536	5	CD750188
128	15	88.2	578	7	AV865990	AV865990 AV865990	c 201	14.4	84.7	538	12	CE007283
129	15	88.2	581	3	BW351213	BW351213 BW351213	c 202	14.4	84.7	539	5	CI268059
130	15	88.2	583	7	AV904914	AV904914 AV904914	c 203	14.4	84.7	541	8	CN467539
131	15	88.2	584	3	BW255065	BW255065 BW255065	c 204	14.4	84.7	541	13	CL585425
132	15	88.2	586	3	BW362639	BW362639 BW362639	c 205	14.4	84.7	542	1	AV435800
133	15	88.2	586	3	BW371870	BW371870 BW371870	c 206	14.4	84.7	546	9	DB359093
134	15	88.2	588	3	BW328828	BW328828 BW328828	c 207	14.4	84.7	548	5	CF101438
135	15	88.2	597	3	BW385928	BW385928 BW385928	c 208	14.4	84.7	548	12	CG810609
136	15	88.2	598	3	BW335229	BW335229 BW335229	c 209	14.4	84.7	549	13	CG738428
137	15	88.2	599	10	DR769559	DR769559 ILLUMIGEN	c 210	14.4	84.7	557	13	CG402242
138	15	88.2	601	3	BW357443	BW357443 BW357443	c 211	14.4	84.7	558	5	CF099984
139	15	88.2	602	3	BW251973	BW251973 BW251973	c 212	14.4	84.7	558	5	CF100002
140	15	88.2	604	7	AV966480	AV966480 AV966480	c 213	14.4	84.7	562	5	CF101525
141	15	88.2	616	3	BW399406	BW399406 BW399406	c 214	14.4	84.7	564	5	CJ158857
142	15	88.2	619	8	CV997966	CV997966 Mdbt60241	c 215	14.4	84.7	573	9	DN606660
143	15	88.2	647	3	BW073840	BW073840 BW073840	c 216	14.4	84.7	581	3	BP352171
144	15	88.2	652	3	BW268878	BW268878 BW268878	c 217	14.4	84.7	582	3	BP287231
145	15	88.2	652	3	BW331827	BW331827 BW331827	c 218	14.4	84.7	582	3	BP353875
146	15	88.2	658	3	BW345318	BW345318 BW345318	c 219	14.4	84.7	582	3	BP354894
147	15	88.2	658	8	CO903450	CO903450 Mddb5022f	c 220	14.4	84.7	583	3	BP287760
148	15	88.2	672	14	DX218546	DX218546 OR Aba009	c 221	14.4	84.7	585	12	CC424344
149	15	88.2	674	3	BW209329	BW209329 BW209329	c 222	14.4	84.7	592	8	CN057494
150	15	88.2	674	5	CK232804	CK232804 ILLUMIGEN	c 223	14.4	84.7	594	14	DX229067
151	15	88.2	715	5	CD871213	CD871213 AZ02.117L	c 224	14.4	84.7	598	5	CD749947
152	15	88.2	782	5	CJ458974	CJ458974 CJ458974	c 225	14.4	84.7	600	5	CI267534
153	15	88.2	805	7	BE727548	BE727548 601560324	c 226	14.4	84.7	600	7	AW732340
154	15	88.2	959	3	BW501533	BW501533 BW501533	c 227	14.4	84.7	605	13	CW083533
155	15	88.2	965	3	BW157642	BW157642 AGENCOURT	c 228	14.4	84.7	606	14	DU807996
156	15	88.2	973	4	BX417613	BX417613 BX417613	c 229	14.4	84.7	607	13	DU255650
157	14.4	84.7	149	11	BZ379755	BZ379755 SALK.1139	c 230	14.4	84.7	608	14	DX001943
158	14.4	84.7	182	3	BQ816777	BQ816777 1030060A1	c 231	14.4	84.7	611	7	BE048187
159	14.4	84.7	186	3	CS095811	CS095811 ie9sa04.b	c 232	14.4	84.7	612	2	BG424419
160	14.4	84.7	201	1	AV039301	AV039301 AV039301	c 233	14.4	84.7	616	2	BI116528
161	14.4	84.7	249	7	BB607340	BB607340 BB607340	c 234	14.4	84.7	622	8	CV960075
162	14.4	84.7	277	12	CG999730	CG999730 ZMWBBb053	c 235	14.4	84.7	628	8	CV893239
163	14.4	84.7	288	12	CG174605	CG174605 PUJDI757B	c 236	14.4	84.7	635	13	CZ347048
164	14.4	84.7	296	1	AI680695	AI680695 tx40c06.x	c 237	14.4	84.7	639	11	BZ351458
165	14.4	84.7	323	7	BF110334	BF110334 7n36f09.x	c 238	14.4	84.7	640	3	BQ868377

239	14.4	84.7	642	14	DX147663	OG Aba018	DX147663	OG Aba018	C 312	14.4	84.7	867	7	BE728024	601561310
240	14.4	84.7	645	13	CL545866	OB_Ba007	CL545866	OB_Ba007	C 313	14.4	84.7	868	3	BQ933100	AGENCY
241	14.4	84.7	648	14	DX110460	OG_Aba013	DX110460	OG_Aba013	C 314	14.4	84.7	869	5	BQ477424	602523647
242	14.4	84.7	650	13	CM311784	104_801_1	CM311784	104_801_1	C 315	14.4	84.7	869	5	CK156247	FGA503718
243	14.4	84.7	652	13	CM311785	104_801_1	CM311785	104_801_1	C 316	14.4	84.7	870	13	BF315359	601902644
244	14.4	84.7	653	14	DX139804	OG_Aba017	DX139804	OG_Aba017	C 317	14.4	84.7	870	13	DU350249	109826819
245	14.4	84.7	655	12	BZ622154	1376612_9	BZ622154	1376612_9	C 318	14.4	84.7	873	2	BG825137	602747925
246	14.4	84.7	659	2	BI180819	3L21F7 LR	BI180819	3L21F7 LR	C 319	14.4	84.7	876	2	BG914111	602805769
247	14.4	84.7	662	14	DU813815	OG_Aba008	DU813815	OG_Aba008	C 320	14.4	84.7	877	10	DM074469	CLPX10198
248	14.4	84.7	673	4	CA175717	SCJES102	CA175717	SCJES102	C 321	14.4	84.7	879	4	BX351545	602437643
249	14.4	84.7	678	8	CO868984	EP_CT_04B	CO868984	EP_CT_04B	C 322	14.4	84.7	879	10	DT802044	127460327
250	14.4	84.7	680	14	AU140676	AU140676	AU140676	AU140676	C 323	14.4	84.7	880	12	CG439468	OG3C281TV
251	14.4	84.7	680	14	DE028888	Branchios	DE028888	Branchios	C 324	14.4	84.7	881	2	BG419034	602446459
252	14.4	84.7	689	8	CM605063	USDA FP 1	CM605063	USDA FP 1	C 325	14.4	84.7	882	13	CZ545108	SRAA-aad5
253	14.4	84.7	692	13	CM087677	104_432_1	CM087677	104_432_1	C 326	14.4	84.7	882	13	CZ545108	SRAA-aad5
254	14.4	84.7	693	8	CV951637	PVpvd_54	CV951637	PVpvd_54	C 327	14.4	84.7	889	5	CF816926	EST694308
255	14.4	84.7	699	8	CV462836	CS_hyp_51	CV462836	CS_hyp_51	C 328	14.4	84.7	891	3	BU150345	AGENCY
256	14.4	84.7	704	8	CO157457	EP_CT_04B	CO157457	EP_CT_04B	C 329	14.4	84.7	891	14	AG899525	Oryza sat
257	14.4	84.7	704	8	CV479128	CS_hyp_41	CV479128	CS_hyp_41	C 330	14.4	84.7	897	12	CC378644	PUPM58TD
258	14.4	84.7	708	13	CM225071	104_562_1	CM225071	104_562_1	C 331	14.4	84.7	898	4	BX363343	601861791
259	14.4	84.7	713	12	CL008944	ZMWBB054	CL008944	ZMWBB054	C 332	14.4	84.7	900	5	CK324731	CNSN01-P-
260	14.4	84.7	714	7	BE566811	6013339409	BE566811	6013339409	C 333	14.4	84.7	902	3	BU508450	AGENCY
261	14.4	84.7	720	8	CO157439	EP_CT_04D	CO157439	EP_CT_04D	C 334	14.4	84.7	903	3	BU150288	AGENCY
262	14.4	84.7	720	14	AG823469	Sus scroF	AG823469	Sus scroF	C 335	14.4	84.7	915	10	DM596961	CGX127-H0
263	14.4	84.7	721	4	CA153440	SCUTR2202	CA153440	SCUTR2202	C 336	14.4	84.7	925	2	BG507145	601861791
264	14.4	84.7	721	8	CO869032	EP_CT_06H	CO869032	EP_CT_06H	C 337	14.4	84.7	927	7	BE737482	601360801
265	14.4	84.7	722	5	CI145507	CI145507	CI145507	CI145507	C 338	14.4	84.7	928	2	BG477718	602521411
266	14.4	84.7	722	8	CO868812	EP_CT_02F	CO868812	EP_CT_02F	C 339	14.4	84.7	945	3	BQ929372	AGENCY
267	14.4	84.7	729	1	AI517687	GH28735_5	AI517687	GH28735_5	C 340	14.4	84.7	948	5	CK164713	FGA504863
268	14.4	84.7	740	12	CC387235	PUEK46TD	CC387235	PUEK46TD	C 341	14.4	84.7	956	8	CK106255	BI125L09
269	14.4	84.7	745	2	BI553178	603197778	BI553178	603197778	C 342	14.4	84.7	960	13	CF918624	EDCA636TR
270	14.4	84.7	747	11	BZ550192	pac81-60	BZ550192	pac81-60	C 343	14.4	84.7	963	7	CF205702	601868478
271	14.4	84.7	750	3	BP875456	BP875456	BP875456	BP875456	C 344	14.4	84.7	967	14	CNS03ACR	Tetraodon
272	14.4	84.7	750	3	BP875456	BP875456	BP875456	BP875456	C 345	14.4	84.7	975	13	CW918670	EDCA660TR
273	14.4	84.7	750	8	CV898729	PBO14E11	CV898729	PBO14E11	C 346	14.4	84.7	977	14	DU768297	ANIW9474
274	14.4	84.7	750	12	CG337214	OG5CY54TC	CG337214	OG5CY54TC	C 347	14.4	84.7	979	14	DU780659	ASXB3327
275	14.4	84.7	754	5	CK183775	EST773090	CK183775	EST773090	C 348	14.4	84.7	979	14	DU792427	APK4559
276	14.4	84.7	754	5	CK183776	UI-H-PH1	CK183776	UI-H-PH1	C 349	14.4	84.7	983	14	CNS079XN	T7 end of
277	14.4	84.7	758	3	BU620252	UI-H-PH1	BU620252	UI-H-PH1	C 350	14.4	84.7	990	14	DU757052	ASNG1333
278	14.4	84.7	762	7	BF573133	602078792	BF573133	602078792	C 351	14.4	84.7	993	14	DU772671	APKG3077
279	14.4	84.7	763	3	BP875457	BP875457	BP875457	BP875457	C 352	14.4	84.7	1018	3	BU172805	AGENCY
280	14.4	84.7	764	13	CZ134645	OA_BBa002	CZ134645	OA_BBa002	C 353	14.4	84.7	1019	7	BF205860	601868263
281	14.4	84.7	766	14	AG823174	Sus scroF	AG823174	Sus scroF	C 354	14.4	84.7	1021	13	CZ995968	184503 TO
282	14.4	84.7	768	14	DX350145	OR_Aba028	DX350145	OR_Aba028	C 355	14.4	84.7	1022	14	DU732827	APK11930
283	14.4	84.7	781	3	BU560126	AGENCY	BU560126	AGENCY	C 356	14.4	84.7	1026	14	DU779809	ASXB2883
284	14.4	84.7	782	2	BG282110	602403248	BG282110	602403248	C 357	14.4	84.7	1026	14	DU742225	ASNC1106
285	14.4	84.7	784	2	BG826004	602403248	BG826004	602403248	C 358	14.4	84.7	1032	2	BG704044	602687178
286	14.4	84.7	785	2	BI149123	602913923	BI149123	602913923	C 359	14.4	84.7	1033	14	DU782699	ASXB560_b
287	14.4	84.7	788	14	AG823039	Sus scroF	AG823039	Sus scroF	C 360	14.4	84.7	1033	14	CNS076PC	AL41326 T3 end of
288	14.4	84.7	791	5	CJ461619	CJ461619	CJ461619	CJ461619	C 361	14.4	84.7	1044	2	BG339349	602437643
289	14.4	84.7	791	5	BE743521	601573023	BE743521	601573023	C 362	14.4	84.7	1058	14	DU793061	APKH4848
290	14.4	84.7	793	3	BP875458	BP875458	BP875458	BP875458	C 363	14.4	84.7	1058	10	DW598302	CGX133-G1
291	14.4	84.7	796	3	BP875458	BP875458	BP875458	BP875458	C 364	14.4	84.7	1092	5	CF585584	AGENCY
292	14.4	84.7	804	10	DM106354	CLRX4120	DM106354	CLRX4120	C 365	14.4	84.7	1096	3	BQ736155	AGENCY
293	14.4	84.7	806	13	DU301652	109844827	DU301652	109844827	C 366	14.4	84.7	1119	9	DF127034	49341108
294	14.4	84.7	812	3	BQ441392	AGENCY	BQ441392	AGENCY	C 367	14.4	84.7	1132	7	BF203527	601866212
295	14.4	84.7	817	12	CC378639	PUPM58TB	CC378639	PUPM58TB	C 368	14.4	84.7	1139	3	BU316693	603852887
296	14.4	84.7	819	7	BE512786	601171548	BE512786	601171548	C 369	14.4	84.7	1148	2	BG293369	602390767
297	14.4	84.7	822	11	BH485953	BGEO807F	BH485953	BGEO807F	C 370	14.4	84.7	1222	7	BE960668	601647982
298	14.4	84.7	830	13	DU166837	OG_Aba003	DU166837	OG_Aba003	C 371	14.4	84.7	1297	1	AL933022	AL933022
299	14.4	84.7	831	2	BG399094	602440763	BG399094	602440763	C 372	14.4	84.7	1314	3	BQ669023	AGENCY
300	14.4	84.7	835	12	CG439463	OG3C281TH	CG439463	OG3C281TH	C 373	14.4	84.7	1354	3	BQ439782	AGENCY
301	14.4	84.7	838	10	DV884551	LB0265.CR	DV884551	LB0265.CR	C 374	14.4	84.7	1404	3	BQ717315	AGENCY
302	14.4	84.7	840	13	CW518619	OR_Ba002	CW518619	OR_Ba002	C 375	14.4	84.7	1451	3	BM925164	AGENCY
303	14.4	84.7	841	14	DX322180	OR_Aba024	DX322180	OR_Aba024	C 376	14.4	84.7	1469	2	BI734931	603355937
304	14.4	84.7	844	12	CC375065	PUPM58TD	CC375065	PUPM58TD	C 377	14.4	84.7	1485	11	AQ317618	RFC11-97
305	14.4	84.7	849	2	BG827957	602752156	BG827957	602752156	C 378	14.4	84.7	1485	11	AQ317618	RFC11-97
306	14.4	84.7	854	7	BE250549	600943259	BE250549	600943259	C 379	14.4	84.7	1485	11	AQ317618	RFC11-97
307	14.4	84.7	855	5	CK812731	Raagc173	CK812731	Raagc173	C 380	14.4	84.7	1485	11	AQ317618	RFC11-97
308	14.4	84.7	855	10	DW106622	CLRX4331	DW106622	CLRX4331	C 381	14.4	84.7	1485	11	AQ317618	RFC11-97
309	14.4	84.7	863	12	BZ679339	PUBFC45TD	BZ679339	PUBFC45TD	C 382	14.4	84.7	1485	11	AQ317618	RFC11-97
310	14.4	84.7	863	14	AG823460	Sus scroF	AG823460	Sus scroF	C 383	14.4	84.7	1485	11	AQ317618	RFC11-97
311	14.4	84.7	867	2	BI761073	603043530	BI761073	603043530	C 384	14.4	84.7	1485	11	AQ317618	RFC11-97

C 385	14	82.4	269	5	CD275181	CL389273	RPCI44_28	458	14	82.4	758	7	BF572155	602076945
C 386	14	82.4	275	13	CL389273	CL389273	RPCI44_28	c 459	14	82.4	759	5	CF870248	tr1c022xb
C 387	14	82.4	283	5	CD276059	CD276059	T143B0242	C 460	14	82.4	764	10	CB907003	tr1c081xb
C 388	14	82.4	293	5	CD273500	CD273500	T143B0143	C 461	14	82.4	767	10	DV555489	DC05024C1
C 389	14	82.4	299	5	CD273199	CD273199	T143B0182	C 462	14	82.4	771	10	DW554022	EST esal
C 390	14	82.4	311	5	CD276119	CD276119	T143B0491	C 463	14	82.4	775	8	CV103889	AGENCOURT-
C 391	14	82.4	314	9	DN435133	DN435133	LIB4217-1	C 464	14	82.4	782	12	CC866160	OGMAX61TV
C 392	14	82.4	320	4	BY482057	BY482057	BY482057	C 465	14	82.4	784	11	BH675266	BOHV473TR
C 393	14	82.4	327	2	BF811797	BF811797	MR2-CI018	C 466	14	82.4	788	4	CB899402	tr1c018x1
C 394	14	82.4	333	8	CN934333	CN934333	000202AVB	C 467	14	82.4	788	13	CL292332	ZMMBB5064
C 395	14	82.4	341	5	CD272899	CD272899	T143B0351	C 468	14	82.4	790	14	CNS038XS	AL233209
C 396	14	82.4	349	9	DN384900	DN384900	LIB3892-0	C 469	14	82.4	792	4	CB897571	Tetradon
C 397	14	82.4	351	2	BM137600	BM137600	WHE0483-0	C 470	14	82.4	797	4	CB896521	tr1c011x1
C 398	14	82.4	412	7	BE184664	BE184664	PM4-HT068	C 471	14	82.4	798	4	CB895258	tr1c001xb
C 400	14	82.4	415	8	CN934520	CN934520	000202AVB	C 472	14	82.4	799	2	BG369432	HVSMEX1002
C 401	14	82.4	430	11	AQ057380	AQ057380	CIT-HSP-2	C 473	14	82.4	801	8	CO810008	AGENCOURT
C 402	14	82.4	448	14	CR154841	CR154841	Forward B	C 474	14	82.4	802	4	CB907070	tr1c081xe
C 403	14	82.4	450	14	CR188559	CR188559	Reverse S	C 475	14	82.4	803	4	CB900908	tr1c024x1
C 404	14	82.4	458	1	AA044686	AA044686	zk75QD09.r	C 476	14	82.4	805	4	CB906771	tr1c080xd
C 405	14	82.4	465	5	CD964283	CD964283	SEB_149_G	C 477	14	82.4	807	2	B1664369	603289932
C 406	14	82.4	481	11	AQ998789	AQ998789	RPCI-23-3	C 478	14	82.4	808	4	CB901470	tr1c026x1
C 407	14	82.4	492	4	BY502850	BY502850	AL815935	C 479	14	82.4	812	4	CB900941	tr1c024xn
C 408	14	82.4	494	1	AL815935	AL815935	AL815935	C 480	14	82.4	812	5	CF870741	tr1c024xn
C 409	14	82.4	506	10	DW996115	DW996115	EST48496	C 481	14	82.4	820	14	CR242684	Reverse S
C 410	14	82.4	512	3	BF744164	BF744164	BP744164	C 482	14	82.4	821	4	CB901058	tr1c025xe
C 411	14	82.4	524	10	T63482	T63482	YC22F01.s1	C 483	14	82.4	828	4	CB900419	tr1c022xb
C 412	14	82.4	536	2	BM140535	BM140535	WHE0484_d	C 484	14	82.4	829	14	AG852780	Oryza sat
C 413	14	82.4	566	8	CN911473	CN911473	021212AEM	C 485	14	82.4	834	14	CT166542	Sus scroF
C 414	14	82.4	567	4	CB511367	CB511367	ssalnwh50	C 486	14	82.4	851	13	CZ189531	OA_BBa011
C 415	14	82.4	568	8	CN883274	CN883274	010813AAS	C 487	14	82.4	896	7	CZ398775	ZMMBF0179
C 416	14	82.4	593	5	CK897513	CK897513	SGP154679	C 488	14	82.4	899	7	BE299550	601892654
C 417	14	82.4	594	13	CZ812008	CZ812008	OC_Ba018	C 489	14	82.4	907	5	CK154423	FGAS03312
C 418	14	82.4	601	3	CZ807597	CZ807597	OC_Ba018	C 490	14	82.4	908	5	CK155055	FGAS03377
C 419	14	82.4	612	8	CN862242	CN862242	000818AAL	C 491	14	82.4	915	3	BU140573	603135884
C 420	14	82.4	616	10	DV094037	DV094037	327-384-4	C 492	14	82.4	915	13	CL206458	ZMMBB0056
C 421	14	82.4	625	8	CN935468	CN935468	000302AVB	C 493	14	82.4	919	6	CNS08PUS	Single re
C 422	14	82.4	636	1	AL808627	AL808627	AL808627	C 494	14	82.4	922	2	BG426462	602492861
C 423	14	82.4	646	4	CA247127	CA247127	SCCFL505	C 495	14	82.4	934	3	BQ958249	AGENCOURT
C 424	14	82.4	654	11	BZ375318	BZ375318	ie38b08.g	C 497	14	82.4	940	10	DV616123	EST121911
C 425	14	82.4	655	4	CB248265	CB248265	UI-M-PD0-	C 498	14	82.4	942	2	BG287829	602384181
C 426	14	82.4	660	12	BZ892074	BZ892074	Hml2_0154	C 499	14	82.4	944	10	DV614053	EST121704
C 427	14	82.4	660	12	CG866148	CG866148	OGMAX611TH	C 500	14	82.4	956	2	BG419094	602446324
C 428	14	82.4	661	3	BM305125	BM305125	BM305125	C 501	14	82.4	971	2	BG750388	602709222
C 429	14	82.4	663	8	CN902740	CN902740	021015ABC	C 502	14	82.4	974	14	DU755197	AGNF551.g
C 430	14	82.4	665	11	B51769	B51769	CIT-HSP-345	C 503	14	82.4	987	2	BG024928	602275870
C 431	14	82.4	672	5	CF883281	CF883281	tr1c088xa	C 504	14	82.4	1017	10	DV039276	CG02_02_E
C 432	14	82.4	688	5	CF875003	CF875003	tr1c037x1	C 505	14	82.4	1075	7	BE618871	601462792
C 433	14	82.4	689	14	AG179835	AG179835	Pan trogl	C 506	14	82.4	1082	9	DR128330	49098865
C 434	14	82.4	691	4	CA369165	CA369165	645542 NC	C 507	14	82.4	1158	7	BF683471	602139658
C 435	14	82.4	693	13	CL749731	CL749731	OR_BBa011	C 508	14	82.4	1212	5	CD517976	AGENCOURT
C 436	14	82.4	696	5	CF868454	CF868454	tr1c015xa	C 509	14	82.4	1243	2	BM455284	AGENCOURT
C 437	14	82.4	699	5	CF872588	CF872588	tr1c031xo	C 510	14	82.4	1287	2	B1545296	603187426
C 438	14	82.4	707	1	AA772995	AA772995	ac48b01.s	C 511	14	82.4	2003	12	CC237578	CH261-191
C 439	14	82.4	709	5	CF879716	CF879716	tr1c081xb	C 512	14	82.4	2391	6	AK017603	Mus muscu
C 440	14	82.4	715	14	CT199313	CT199313	Sus scroF	C 513	14	82.4	3427	6	AK017603	Mus muscu
C 441	14	82.4	727	7	BE888245	BE888245	601511766	C 514	13.8	81.2	112	7	BF170493	Pongo pyg
C 442	14	82.4	728	4	CB909372	CB909372	tr1c088xa	C 515	13.8	81.2	116	8	CD265744	FCLO940 M
C 443	14	82.4	728	7	BE383439	BE383439	601297254	C 516	13.8	81.2	118	2	BF802075	170004247
C 444	14	82.4	729	5	CF879767	CF879767	tr1c081xe	C 517	13.8	81.2	126	14	CR170997	CM1-CI009
C 445	14	82.4	729	8	CO435008	CO435008	595_CAS-0	C 518	13.8	81.2	136	5	CK007466	28658rsic
C 446	14	82.4	732	5	CF869267	CF869267	tr1c018x1	C 519	13.8	81.2	136	5	CF839592	PSHH002xN
C 447	14	82.4	736	5	CF867533	CF867533	tr1c011x1	C 520	13.8	81.2	149	8	CV933694	PMprcm_36
C 448	14	82.4	739	14	CR003341	CR003341	Reverse S	C 521	13.8	81.2	152	6	AK192906	Mus muscu
C 449	14	82.4	742	5	CF865366	CF865366	tr1c001xb	C 522	13.8	81.2	184	10	T08267	EST06158 In
C 450	14	82.4	742	5	CF866542	CF866542	tr1c007xb	C 523	13.8	81.2	197	2	BF874254	IL3-ET011
C 451	14	82.4	744	4	CB904261	CB904261	tr1c037x1	C 524	13.8	81.2	203	10	T55169	Yb45g10.r1
C 452	14	82.4	750	5	CF870710	CF870710	tr1c024x1	C 525	13.8	81.2	204	2	BG953416	PM3-CT064
C 453	14	82.4	752	5	CF870852	CF870852	tr1c025xe	C 526	13.8	81.2	206	9	DA695360	DA695360
C 454	14	82.4	752	4	CB989533	CB989533	tr1c015xa	C 527	13.8	81.2	211	3	BM764193	K-EST0045
C 455	14	82.4	752	5	CF871246	CF871246	tr1c026x1	C 528	13.8	81.2	212	9	D56329	HUM424F08B
C 456	14	82.4	752	5	CF879339	CF879339	tr1c080xd	C 529	13.8	81.2	212	13	CW847798	GT15098.D
C 457	14	82.4	754	4	CB902833	CB902833	tr1c031xo	C 530	13.8	81.2	214	7	BE078145	CM0-BT061

C 531	13.8	81.2	216	11	B15103	B15103 344A2-TV CI	604	13.8	81.2	317	9	DA005806	DA005806
C 532	13.8	81.2	217	1	AU280093	AU280093 AU280093	C 605	13.8	81.2	320	9	DB028064	DB028064
C 533	13.8	81.2	218	9	DA157320	DA157320 DA157320	C 606	13.8	81.2	321	7	BE655458	BE655458
C 534	13.8	81.2	219	1	AA953015	AA953015 on67h03.s	C 607	13.8	81.2	322	3	EG890244	EG890244
C 535	13.8	81.2	220	7	BB563534	BB563534 BB563534	C 608	13.8	81.2	323	4	BY331549	BY331549
C 536	13.8	81.2	221	7	B1029709	B1029709 MR4-MT025	C 609	13.8	81.2	324	4	BY346686	BY346686
C 537	13.8	81.2	222	3	CJ183774	CJ183774 CJ183774	C 610	13.8	81.2	325	8	CN408195	CN408195
C 538	13.8	81.2	223	5	DA151756	DA151756 DA151756	C 611	13.8	81.2	326	2	BF947367	BF947367
C 539	13.8	81.2	224	1	AV262236	AV262236 AV262236	C 612	13.8	81.2	327	4	BY330281	BY330281
C 540	13.8	81.2	225	1	AA369613	AA369613 EST81047	C 613	13.8	81.2	328	4	BY377409	BY377409
C 541	13.8	81.2	226	7	BE860151	BE860151 UI-M-A10-	C 614	13.8	81.2	329	7	AW163647	AW163647
C 542	13.8	81.2	227	13	CW046889	CW046889 104_284_1	C 615	13.8	81.2	330	7	BB779343	BB779343
C 543	13.8	81.2	228	11	BH755995	BH755995 SALK_0526	C 616	13.8	81.2	331	11	AQ843783	AQ843783
C 544	13.8	81.2	229	11	BH677240	BH677240 BOHWP43TF	C 617	13.8	81.2	332	1	AJ788883	AJ788883
C 545	13.8	81.2	230	4	C18116	C18116 C18116 Huma	C 618	13.8	81.2	333	10	H74254	H74254
C 546	13.8	81.2	231	7	BB638571	BB638571 BB638571	C 619	13.8	81.2	334	10	R97710	R97710
C 547	13.8	81.2	232	13	CZ105519	CZ105519 OM_Ba011	C 620	13.8	81.2	335	9	D15565	D15565
C 548	13.8	81.2	233	4	BY596518	BY596518 BY596518	C 621	13.8	81.2	336	4	BY331704	BY331704
C 549	13.8	81.2	234	8	CN198913	CN198913 TgESTv16	C 622	13.8	81.2	337	7	AW426541	AW426541
C 550	13.8	81.2	235	11	BZ154263	BZ154263 CH230-311	C 623	13.8	81.2	338	2	BM146327	BM146327
C 551	13.8	81.2	236	9	DA846603	DA846603 DA846603	C 624	13.8	81.2	339	4	BY129872	BY129872
C 552	13.8	81.2	237	4	BY232080	BY232080 BY232080	C 625	13.8	81.2	340	10	T68314	T68314
C 553	13.8	81.2	238	11	BH865576	BH865576 SALK_0989	C 626	13.8	81.2	341	11	AQ906499	AQ906499
C 554	13.8	81.2	239	7	BE357618	BE357618 BB357618	C 627	13.8	81.2	342	4	BY780631	BY780631
C 555	13.8	81.2	240	7	BE206222	BE206222 ba93f06.x	C 628	13.8	81.2	343	4	BY141469	BY141469
C 556	13.8	81.2	241	7	BE013658	BE013658 125082.MA	C 629	13.8	81.2	344	4	BY325811	BY325811
C 557	13.8	81.2	242	5	CF060502	CF060502 QCT10C04.	C 630	13.8	81.2	345	4	BY351793	BY351793
C 558	13.8	81.2	243	5	CK460335	CK460335 930604.MA	C 631	13.8	81.2	346	4	BY353335	BY353335
C 559	13.8	81.2	244	5	CK461038	CK461038 931372.MA	C 632	13.8	81.2	347	4	BY794050	BY794050
C 560	13.8	81.2	245	9	DN102908	DN102908 1097918.M	C 633	13.8	81.2	348	7	BB846887	BB846887
C 561	13.8	81.2	246	12	CC050731	CC050731 SALK_0009	C 634	13.8	81.2	349	4	CA545246	CA545246
C 562	13.8	81.2	247	6	AK191737	AK191737 Mus_musc	C 635	13.8	81.2	350	7	AW427501	AW427501
C 563	13.8	81.2	248	7	DA255846	DA255846 DA255846	C 636	13.8	81.2	351	9	DA632005	DA632005
C 564	13.8	81.2	249	10	DT656367	DT656367 P3rin.UAO	C 637	13.8	81.2	352	4	BY352913	BY352913
C 565	13.8	81.2	250	5	CF062956	CF062956 QCU15G04.	C 638	13.8	81.2	353	4	BY778435	BY778435
C 566	13.8	81.2	251	7	BB577069	BB577069 BB577069	C 639	13.8	81.2	354	8	CR520514	CR520514
C 567	13.8	81.2	252	7	BS554661	BS554661 BS554661	C 640	13.8	81.2	355	11	AQ850907	AQ850907
C 568	13.8	81.2	253	1	AI653718	AI653718 wb36e10.x	C 641	13.8	81.2	356	3	BU077708	BU077708
C 569	13.8	81.2	254	2	EG712966	EG712966 pglrn.pk0	C 642	13.8	81.2	357	7	AW583725	AW583725
C 570	13.8	81.2	255	2	BG433455	BG433455 602507043	C 643	13.8	81.2	358	4	BY331570	BY331570
C 571	13.8	81.2	256	2	BG542469	BG542469 602569547	C 644	13.8	81.2	359	3	BM764060	BM764060
C 572	13.8	81.2	257	8	CN653697	CN653697 TgESTvtn4	C 645	13.8	81.2	360	4	BY791775	BY791775
C 573	13.8	81.2	258	12	CE028388	CE028388 tigr-gss-	C 646	13.8	81.2	361	9	DN539044	DN539044
C 574	13.8	81.2	259	3	BM844454	BM844454 K-EST0122	C 647	13.8	81.2	362	4	BY101143	BY101143
C 575	13.8	81.2	260	14	DU993173	DU993173 OG_Aba011	C 648	13.8	81.2	363	10	DY205550	DY205550
C 576	13.8	81.2	261	7	AW054731	AW054731 wz97906.x	C 649	13.8	81.2	364	4	BY334271	BY334271
C 577	13.8	81.2	262	13	CL432141	CL432141 ZMWBBb044	C 650	13.8	81.2	365	4	CB693320	CB693320
C 578	13.8	81.2	263	7	AA112399	AA112399 zm27C04.r	C 651	13.8	81.2	366	4	BY353055	BY353055
C 579	13.8	81.2	264	7	BB346865	BB346865 BB346865	C 652	13.8	81.2	367	4	BY305334	BY305334
C 580	13.8	81.2	265	2	BI041496	BI041496 MR4-NT014	C 653	13.8	81.2	368	8	CV606155	CV606155
C 581	13.8	81.2	266	9	DB171839	DB171839 DB171839	C 654	13.8	81.2	369	4	BY235476	BY235476
C 582	13.8	81.2	267	9	DA363581	DA363581 DA363581	C 655	13.8	81.2	370	4	BY318363	BY318363
C 583	13.8	81.2	268	1	AU099791	AU099791 AU099791	C 656	13.8	81.2	371	4	BY785879	BY785879
C 584	13.8	81.2	269	3	DA091478	DA091478 DA091478	C 657	13.8	81.2	372	8	CN486637	CN486637
C 585	13.8	81.2	270	7	AW026964	AW026964 wv64b03.x	C 658	13.8	81.2	373	10	H52326	H52326
C 586	13.8	81.2	271	2	BG904622	BG904622 Talx1133F	C 659	13.8	81.2	374	4	BY346520	BY346520
C 587	13.8	81.2	272	3	BI307196	BI307196 PTS0535	C 660	13.8	81.2	375	4	BY793509	BY793509
C 588	13.8	81.2	273	9	DA265182	DA265182 DA265182	C 661	13.8	81.2	376	1	AA969019	AA969019
C 589	13.8	81.2	274	9	DA363454	DA363454 DA363454	C 662	13.8	81.2	377	3	BY325120	BY325120
C 590	13.8	81.2	275	11	AF524463	AF524463 AF524463	C 663	13.8	81.2	378	4	BY134403	BY134403
C 591	13.8	81.2	276	3	CL408259	CL408259 ZMWBBb027	C 664	13.8	81.2	379	4	BY194686	BY194686
C 592	13.8	81.2	277	1	AA407301	AA407301 EST02298	C 665	13.8	81.2	380	5	CJ010631	CJ010631
C 593	13.8	81.2	278	4	CB698158	CB698158 AMGNNUC:M	C 666	13.8	81.2	381	8	CO944622	CO944622
C 594	13.8	81.2	279	9	DB168544	DB168544 DB168544	C 667	13.8	81.2	382	10	DR798316	DR798316
C 595	13.8	81.2	280	3	BM825895	BM825895 K-EST0097	C 668	13.8	81.2	383	4	A1200735	A1200735
C 596	13.8	81.2	281	7	BE242314	BE242314 TCAAP1T20	C 669	13.8	81.2	384	1	BY304092	BY304092
C 597	13.8	81.2	282	10	DY062712	DY062712 001013BEM	C 670	13.8	81.2	385	4	BY776324	BY776324
C 598	13.8	81.2	283	13	CL227862	CL227862 ZMWBBc002	C 671	13.8	81.2	386	9	DB017588	DB017588
C 599	13.8	81.2	284	9	DA915859	DA915859 DA915859	C 672	13.8	81.2	387	9	DN277274	DN277274
C 600	13.8	81.2	285	1	AM071742	AM071742 AM071742	C 673	13.8	81.2	388	3	BY333383	BY333383
C 601	13.8	81.2	286	1	AV639598	AV639598 AV639598	C 674	13.8	81.2	389	1	A1207876	A1207876
C 602	13.8	81.2	287	4	BY130941	BY130941 BY130941	C 675	13.8	81.2	390	4	BY339667	BY339667
C 603	13.8	81.2	288	4	BY347564	BY347564 BY347564	C 676	13.8	81.2	391	4	CB813213	CB813213

677	13.8	81.2	357	7	B8872913	B8872913	B8872913	C 750	13.8	81.2	394	9	DA662409	DA662409
678	13.8	81.2	358	4	BY214746	BY214746	BY214746	C 751	13.8	81.2	394	10	DM988250	DM988250
679	13.8	81.2	358	7	B8864107	UI-M-BH1-	BE864107	C 752	13.8	81.2	395	4	CB775816	CB775816
680	13.8	81.2	358	9	DA000576	DA000576	DA000576	C 753	13.8	81.2	395	13	CBW355043	CBW355043
681	13.8	81.2	358	9	DB014769	DB014769	DB014769	C 754	13.8	81.2	396	1	AA056385	AA056385
682	13.8	81.2	359	4	BY301412	BY301412	BY301412	C 755	13.8	81.2	396	4	C19178	C19178
683	13.8	81.2	359	4	BY335207	BY335207	BY335207	C 756	13.8	81.2	396	5	CJ282128	CJ282128
684	13.8	81.2	360	4	CU1831	CU1831	CU1831	C 757	13.8	81.2	396	12	BZ2915094	BZ2915094
685	13.8	81.2	360	7	BF523326	UI-R-GO-u	BF523326	C 758	13.8	81.2	397	4	BY229746	BY229746
686	13.8	81.2	360	8	CV676217	ucg1h08.k	CV676217	C 759	13.8	81.2	397	4	BY533326	BY533326
687	13.8	81.2	361	4	C82576	C82576	C82576	C 760	13.8	81.2	398	5	CJ170137	CJ170137
688	13.8	81.2	361	4	C83432	C83432	C83432	C 761	13.8	81.2	398	3	BM845706	BM845706
689	13.8	81.2	362	1	AV393579	AV393579	AV393579	C 762	13.8	81.2	399	4	BY018720	BY018720
690	13.8	81.2	362	4	BY232135	BY232135	BY232135	C 763	13.8	81.2	399	4	BY100198	BY100198
691	13.8	81.2	362	5	CU037407	CU037407	CU037407	C 764	13.8	81.2	399	5	CJ086377	CJ086377
692	13.8	81.2	362	8	CX201758	MN503270	CX201758	C 765	13.8	81.2	399	13	CZ1703264	CZ1703264
693	13.8	81.2	363	2	BG158871	RHI22.43	BG158871	C 766	13.8	81.2	400	1	AL363918	AL363918
694	13.8	81.2	364	3	B8648818	B8648818	B8648818	C 767	13.8	81.2	400	5	CD627486	CD627486
695	13.8	81.2	365	3	BP097748	BP097748	BP097748	C 768	13.8	81.2	400	5	CJ172599	CJ172599
696	13.8	81.2	365	4	BY215352	BY215352	BY215352	C 769	13.8	81.2	400	6	AK200415	AK200415
697	13.8	81.2	365	5	CN292358	C1292358	CN292358	C 770	13.8	81.2	400	10	DR393071	DR393071
698	13.8	81.2	365	9	DN877900	nael7a12.	DN877900	C 771	13.8	81.2	401	1	AI929516	AI929516
699	13.8	81.2	366	3	BY083781	BY083781	BY083781	C 772	13.8	81.2	401	5	CJ081947	CJ081947
700	13.8	81.2	366	4	BY199956	BY199956	BY199956	C 773	13.8	81.2	402	3	BM857359	BM857359
701	13.8	81.2	366	4	BY442246	BY442246	BY442246	C 774	13.8	81.2	402	4	CB769261	CB769261
702	13.8	81.2	367	5	CJ737180	CJ737180	CJ737180	C 775	13.8	81.2	402	7	BB733381	BB733381
703	13.8	81.2	367	9	DA918467	DA918467	DA918467	C 776	13.8	81.2	402	9	DA942173	DA942173
704	13.8	81.2	368	5	CJ179816	CJ179816	CJ179816	C 777	13.8	81.2	405	1	AI931255	AI931255
705	13.8	81.2	368	7	AW044227	wy68g06.x	AW044227	C 778	13.8	81.2	405	7	BB230373	BB230373
706	13.8	81.2	370	9	DN615997	SPM21C11	DN615997	C 779	13.8	81.2	406	1	AA968770	AA968770
707	13.8	81.2	371	4	BY313568	BY313568	BY313568	C 780	13.8	81.2	407	2	B1248079	B1248079
708	13.8	81.2	372	4	BY320250	BY320250	BY320250	C 781	13.8	81.2	407	4	BX540890	BX540890
709	13.8	81.2	373	4	BY036686	BY036686	BY036686	C 782	13.8	81.2	407	4	CB771836	CB771836
710	13.8	81.2	374	10	DV227798	EST-AR161	DV227798	C 783	13.8	81.2	409	1	AI005645	AI005645
711	13.8	81.2	374	11	AQ528337	RPCI-11-3	AQ528337	C 784	13.8	81.2	409	1	AV598895	AV598895
712	13.8	81.2	374	12	CQ597590	ZMWBBC040	CQ597590	C 785	13.8	81.2	409	1	AV660393	AV660393
713	13.8	81.2	375	4	CB691544	AMGNNUC:S	CB691544	C 786	13.8	81.2	409	2	BI537426	BI537426
714	13.8	81.2	375	7	AW294667	UI-H-BM0-	AW294667	C 787	13.8	81.2	409	2	BM032222	BM032222
715	13.8	81.2	378	4	BY767298	BY767298	BY767298	C 788	13.8	81.2	409	3	BM836277	BM836277
716	13.8	81.2	379	4	BY133379	BY133379	BY133379	C 789	13.8	81.2	409	4	BY018868	BY018868
717	13.8	81.2	379	7	B8842586	B8842586	B8842586	C 790	13.8	81.2	409	4	BY268271	BY268271
718	13.8	81.2	379	9	DA665725	DA665725	DA665725	C 791	13.8	81.2	409	4	CB804829	CB804829
719	13.8	81.2	379	13	CL286180	ZMWBD063	CL286180	C 792	13.8	81.2	409	4	CD057143	CD057143
720	13.8	81.2	379	14	AG952911	Drcosophil	AG952911	C 793	13.8	81.2	409	5	CJ034596	CJ034596
721	13.8	81.2	380	12	CQ987804	CH240.138	CQ987804	C 794	13.8	81.2	409	5	CJ178191	CJ178191
722	13.8	81.2	381	4	BY099696	BY099696	BY099696	C 795	13.8	81.2	409	13	CM463386	CM463386
723	13.8	81.2	381	4	CB809847	AMGNNUC:T	CB809847	C 796	13.8	81.2	410	4	BY289366	BY289366
724	13.8	81.2	382	11	AZ327355	1M0050E12	AZ327355	C 797	13.8	81.2	410	4	BY536810	BY536810
725	13.8	81.2	382	14	DU589822	OO_Ba008	DU589822	C 798	13.8	81.2	410	9	DA869586	DA869586
726	13.8	81.2	383	2	BF873291	CM0-ET012	BF873291	C 799	13.8	81.2	411	7	AW464822	AW464822
727	13.8	81.2	383	7	B8856822	BB856822	B8856822	C 800	13.8	81.2	411	9	BB740260	BB740260
728	13.8	81.2	383	8	CV384989	QV1-AN005	CV384989	C 801	13.8	81.2	411	9	DB024399	DB024399
729	13.8	81.2	384	4	BY768935	BY768935	BY768935	C 802	13.8	81.2	412	11	BH881931	BH881931
730	13.8	81.2	384	10	B67023	me29c12.r1	B67023	C 803	13.8	81.2	412	4	CB121146	CB121146
731	13.8	81.2	385	4	BY773170	BY773170	BY773170	C 804	13.8	81.2	414	5	CJ177984	CJ177984
732	13.8	81.2	385	5	CJ198602	CJ198602	CJ198602	C 805	13.8	81.2	414	9	DA547688	DA547688
733	13.8	81.2	386	1	A1Z70033	qt91b05.x	A1Z70033	C 806	13.8	81.2	414	9	DB183371	DB183371
734	13.8	81.2	386	8	CR549467	DKFZp468H	CR549467	C 807	13.8	81.2	416	6	AK186576	AK186576
735	13.8	81.2	387	1	A1854554	UI-M-BH0-	A1854554	C 808	13.8	81.2	416	7	BB846851	BB846851
736	13.8	81.2	387	2	BM194527	TCAAP2011	BM194527	C 809	13.8	81.2	417	4	BY3113819	BY3113819
737	13.8	81.2	388	4	CA211744	scr1AD110	CA211744	C 810	13.8	81.2	417	5	CJ279008	CJ279008
738	13.8	81.2	389	1	AA667014	vr86f01.s	AA667014	C 811	13.8	81.2	418	5	CJ171530	CJ171530
739	13.8	81.2	389	8	CV356790	mr432d12.r1	CV356790	C 812	13.8	81.2	418	7	BE009730	BE009730
740	13.8	81.2	389	10	WB3889	MEF32d12.r1	WB3889	C 813	13.8	81.2	418	12	BZ782109	BZ782109
741	13.8	81.2	390	4	BY520734	BY520734	BY520734	C 814	13.8	81.2	419	2	BM194021	BM194021
742	13.8	81.2	390	9	DA176390	DA176390	DA176390	C 815	13.8	81.2	419	7	BB687745	BB687745
743	13.8	81.2	390	9	DN274429	DN274429	DN274429	C 816	13.8	81.2	419	13	DU120189	DU120189
744	13.8	81.2	391	1	A1854558	UI-M-BH0-	A1854558	C 817	13.8	81.2	420	7	BE375637	BE375637
745	13.8	81.2	391	13	CMW13657	104.517.1	CMW13657	C 818	13.8	81.2	421	1	AA628180	AA628180
746	13.8	81.2	393	3	BQ503976	EST13206-	BQ503976	C 819	13.8	81.2	421	7	AW285090	AW285090
747	13.8	81.2	393	5	CJ170170	CJ170170	CJ170170	C 820	13.8	81.2	421	10	R98493	R98493
748	13.8	81.2	393	5	CJ170355	CJ170355	CJ170355	C 821	13.8	81.2	422	1	AA737790	AA737790
749	13.8	81.2	393	5	CJ179071	CJ179071	CJ179071	C 822	13.8	81.2	424	2	BG636712	BG636712

823	13.8	81.2	425	4	BY547337	BY547337	BY547337	BY547337	C 896	13.8	81.2	454	9	DA552241	DA552241
824	13.8	81.2	425	4	BY633000	BY633000	BY633000	BY633000	C 897	13.8	81.2	455	3	DA537667	DA537667
825	13.8	81.2	425	12	CC888292	CC888292	SALK_1516	CC888292	C 898	13.8	81.2	456	3	BM836400	K-EST0112
826	13.8	81.2	426	4	BX475460	DXFPZP686J	BX475460	DXFPZP686J	C 899	13.8	81.2	456	9	DA083628	DA083628
827	13.8	81.2	426	7	BB728801	BB728801	BB728801	BB728801	C 900	13.8	81.2	457	7	BE856791	7E68H02.X
828	13.8	81.2	426	9	DA994360	DA994360	DA994360	DA994360	C 901	13.8	81.2	458	4	BY571085	BY571085
829	13.8	81.2	429	4	BY277616	BY277616	BY277616	BY277616	C 902	13.8	81.2	458	4	DA987071	DA987071
830	13.8	81.2	429	4	BY288598	BY288598	BY288598	BY288598	C 903	13.8	81.2	458	9	DB251063	DB251063
831	13.8	81.2	429	4	BY632109	BY632109	BY632109	BY632109	C 904	13.8	81.2	459	4	BY599624	BY599624
832	13.8	81.2	429	5	CJ176724	CJ176724	CJ176724	CJ176724	C 905	13.8	81.2	459	7	AW465372	BP230019A
833	13.8	81.2	429	7	AW249026	AW249026	2820760.5	AW249026	C 906	13.8	81.2	459	9	DA087387	DA087387
834	13.8	81.2	430	5	CJ165752	CJ165752	CJ165752	CJ165752	C 907	13.8	81.2	460	4	CB737756	ANGNUUC:N
835	13.8	81.2	431	1	AI155262	AI155262	u87f066.r	AI155262	C 908	13.8	81.2	460	5	CF339261	RCL1--04-
836	13.8	81.2	431	3	BM750982	BM750982	K-EST0026	BM750982	C 909	13.8	81.2	460	9	DA4894	RIC514713A
837	13.8	81.2	431	4	BY939466	BY939466	BY939466	BY939466	C 910	13.8	81.2	462	2	B1534940	B1534940
838	13.8	81.2	431	7	BB727333	BB727333	BB727333	BB727333	C 911	13.8	81.2	462	14	CR9123987	Sus scrofa
839	13.8	81.2	431	14	AG976917	AG976917	Drosophila	AG976917	C 912	13.8	81.2	463	4	BY251955	BY251955
840	13.8	81.2	432	4	BY167535	BY167535	BY167535	BY167535	C 913	13.8	81.2	463	5	CF034830	CF034830
841	13.8	81.2	432	4	BY211827	BY211827	BY211827	BY211827	C 914	13.8	81.2	463	8	BE390981	UI-M-BH3
842	13.8	81.2	432	4	BY539396	BY539396	BY539396	BY539396	C 915	13.8	81.2	464	8	CV514070	TgESTzyr0
843	13.8	81.2	432	10	DV471104	DV471104	5252_C08	DV471104	C 916	13.8	81.2	466	4	BA432512	BA432512
844	13.8	81.2	432	10	T56251	T56251	y632h01.s1	T56251	C 917	13.8	81.2	466	5	CF039208	QC33d10.
845	13.8	81.2	433	8	CN199470	CN199470	TgESTzy17	CN199470	C 918	13.8	81.2	466	7	BB861165	BB861165
846	13.8	81.2	434	1	AM055603	AM055603	AM055603	AM055603	C 919	13.8	81.2	466	8	CN336468	CN336468
847	13.8	81.2	434	5	AA432216	AA432216	V635g12.r	AA432216	C 920	13.8	81.2	467	5	CF038757	QC28C01.
848	13.8	81.2	434	5	BE267536	BE267536	BE267536	BE267536	C 921	13.8	81.2	467	9	DA058037	DA058037
849	13.8	81.2	434	8	BM788135	BM788135	BM788135	BM788135	C 922	13.8	81.2	468	5	CF036902	QC338e11.
850	13.8	81.2	434	8	CN154135	CN154135	941322_MA	CN154135	C 923	13.8	81.2	468	7	BF599892	264279_MA
851	13.8	81.2	434	9	CA606283	mda94f11.	CA606283	mda94f11.	C 924	13.8	81.2	468	14	AJ594973	Arabiadops
852	13.8	81.2	435	1	AM055603	AM055603	AM055603	AM055603	C 925	13.8	81.2	469	10	NS9391	yv51b03.s1
853	13.8	81.2	435	3	BQ293948	BQ293948	1091024B0	BQ293948	C 926	13.8	81.2	469	14	DX197498	DX197498
854	13.8	81.2	435	7	AW124617	AW124617	UI-M-BH2.	AW124617	C 927	13.8	81.2	470	1	AI788270	UI788270
855	13.8	81.2	435	7	BB727209	BB727209	BB727209	BB727209	C 928	13.8	81.2	470	3	BQ559278	BQ559278
856	13.8	81.2	436	5	CD704090	EST20617	CD704090	EST20617	C 929	13.8	81.2	470	14	CT445009	Sus scrofa
857	13.8	81.2	437	2	BI910678	603070445	BI910678	603070445	C 930	13.8	81.2	471	3	BQ270895	BQ270895
858	13.8	81.2	437	14	DX230152	DX230152	OR_Aba011	DX230152	C 931	13.8	81.2	471	3	BY595785	BY595785
859	13.8	81.2	438	5	CI308472	CI308472	CI308472	CI308472	C 932	13.8	81.2	471	7	BE265817	BE265817
860	13.8	81.2	438	9	DB267536	DB267536	DB267536	DB267536	C 933	13.8	81.2	471	9	DA055510	DA055510
861	13.8	81.2	440	4	BY536244	BY536244	BY536244	BY536244	C 934	13.8	81.2	471	9	DA593065	DA593065
862	13.8	81.2	440	5	CI163472	CJ163472	CJ163472	CJ163472	C 935	13.8	81.2	471	9	DA732967	DA732967
863	13.8	81.2	440	9	DA212571	DA212571	DA212571	DA212571	C 936	13.8	81.2	471	9	DN616197	DN616197
864	13.8	81.2	441	12	CG453904	CG453904	OCTBFlITV	CG453904	C 937	13.8	81.2	471	11	AQ937430	AQ937430
865	13.8	81.2	442	4	BY001867	BY001867	BY001867	BY001867	C 938	13.8	81.2	472	5	BM257486	BM257486
866	13.8	81.2	443	5	CJ181464	CJ181464	CJ181464	CJ181464	C 939	13.8	81.2	472	5	CJ176784	CJ176784
867	13.8	81.2	443	7	AW050191	UI-M-BH1-	AW050191	UI-M-BH1-	C 940	13.8	81.2	472	5	CJ286274	CJ286274
868	13.8	81.2	444	2	BF726786	bv12802.y	BF726786	bv12802.y	C 941	13.8	81.2	473	1	AI822046	AI822046
869	13.8	81.2	444	4	CA829662	3529_1.7	CA829662	3529_1.7	C 942	13.8	81.2	473	2	BI218822	BI218822
870	13.8	81.2	444	4	CB086801	h390f04.g	CB086801	h390f04.g	C 943	13.8	81.2	473	4	BY562843	BY562843
871	13.8	81.2	444	13	CL959093	OE1FCC002	CL959093	OE1FCC002	C 944	13.8	81.2	473	4	AW083037	AW083037
872	13.8	81.2	446	4	BY371893	BY371893	BY371893	BY371893	C 945	13.8	81.2	474	9	DA144139	DA144139
873	13.8	81.2	446	5	CI180630	CI180630	CI180630	CI180630	C 946	13.8	81.2	474	9	DA551493	DA551493
874	13.8	81.2	446	12	CE312972	CE312972	tigr-ges-	CE312972	C 947	13.8	81.2	475	2	DB031630	DB031630
875	13.8	81.2	447	5	CJ163613	CJ163613	CJ163613	CJ163613	C 948	13.8	81.2	476	2	BG002276	BG002276
876	13.8	81.2	448	4	CA169065	SCJFSB101	CA169065	SCJFSB101	C 949	13.8	81.2	476	4	CA391293	CA391293
877	13.8	81.2	448	8	CX118286	EX1114D03	CX118286	EX1114D03	C 950	13.8	81.2	476	7	BB857905	BB857905
878	13.8	81.2	448	9	DA683715	DA683715	DA683715	DA683715	C 951	13.8	81.2	476	11	AQ885533	AQ885533
879	13.8	81.2	449	9	DA520844	DA520844	DA520844	DA520844	C 952	13.8	81.2	477	1	AA877930	AA877930
880	13.8	81.2	449	11	AQ172654	HS_3194_A	AQ172654	HS_3194_A	C 953	13.8	81.2	477	10	DV309284	DV309284
881	13.8	81.2	450	4	BY555917	BY555917	BY555917	BY555917	C 954	13.8	81.2	478	1	AA614748	AA614748
882	13.8	81.2	450	4	CA528199	CA528199	8064-01_M	CA528199	C 955	13.8	81.2	478	4	BY640393	BY640393
883	13.8	81.2	450	9	DA825138	DA825138	DA825138	DA825138	C 956	13.8	81.2	479	7	AV937692	AV937692
884	13.8	81.2	450	11	BZ191042	BZ191042	CH230-515	BZ191042	C 957	13.8	81.2	479	12	CJ190604	CJ190604
885	13.8	81.2	451	7	AW476719	u8r1h01.y	AW476719	u8r1h01.y	C 958	13.8	81.2	480	5	CJ149817	CJ149817
886	13.8	81.2	451	7	BB726189	BB726189	BB726189	BB726189	C 959	13.8	81.2	480	5	CJ154704	CJ154704
887	13.8	81.2	452	5	CJ173750	CJ173750	CJ173750	CJ173750	C 960	13.8	81.2	480	9	DA525291	DA525291
888	13.8	81.2	452	9	DA702784	DA702784	DA702784	DA702784	C 961	13.8	81.2	481	5	CJ176215	CJ176215
889	13.8	81.2	453	7	BB776704	BB776704	BB776704	BB776704	C 962	13.8	81.2	481	5	CJ235283	CJ235283
890	13.8	81.2	453	9	DB221655	DB221655	DB221655	DB221655	C 963	13.8	81.2	481	7	BE976340	BE976340
891	13.8	81.2	453	11	AQ687627	nbxb0075E	AQ687627	nbxb0075E	C 964	13.8	81.2	481	9	DA270601	DA270601
892	13.8	81.2	454	4	BY589851	BY589851	BY589851	BY589851	C 965	13.8	81.2	481	9	DA596265	DA596265
893	13.8	81.2	454	5	CJ739092	CJ739092	CJ739092	CJ739092	C 966	13.8	81.2	481	9	DA701978	DA701978
894	13.8	81.2	454	7	BB815206	BB815206	BB815206	BB815206	C 967	13.8	81.2	483	7	AW786192	AW786192
895	13.8	81.2	454	9	DA080055	DA080055	DA080055	DA080055	C 968	13.8	81.2	483	9	DA046437	DA046437

```
c 969      13.8      81.2      483      9      DN278604      1157927 M
c 970      13.8      81.2      485      7      AW762768      ur58904.y
c 971      13.8      81.2      485      9      DA083372      DA083372
c 972      13.8      81.2      485      9      DA836376      DA836376
c 973      13.8      81.2      485      9      DB129653      DB129653
c 974      13.8      81.2      485      14     DU657459      OG_Aba004
c 975      13.8      81.2      486      9      DA531673      DA531673
c 976      13.8      81.2      487      1      AA406139      zu20c01.s
c 977      13.8      81.2      487      4      BY596343      BY596343
c 978      13.8      81.2      487      7      BF563873      UI-R-BT1-
c 979      13.8      81.2      487      9      DA075391      DA075391
c 980      13.8      81.2      487      9      DA079145      DA079145
c 981      13.8      81.2      487      9      DA457837      DA457837
c 982      13.8      81.2      487      9      DA472472      DA472472
c 983      13.8      81.2      487      9      DA692013      DA692013
c 984      13.8      81.2      488      1      AU189877      AU189877
c 985      13.8      81.2      488      4      BX344422      BX344422
c 986      13.8      81.2      488      14     CENS00PT0     AL084930      Arabidops
c 987      13.8      81.2      489      9      DA121025      DA121025
c 988      13.8      81.2      489      9      DA303276      DA303276
c 989      13.8      81.2      489      9      DA358117      DA358117
c 990      13.8      81.2      489      9      DA532761      DA532761
c 991      13.8      81.2      490      2      BM290225      EST576759
c 992      13.8      81.2      490      3      BP200251      BP200251
c 993      13.8      81.2      490      4      CB725836      AMGNNUC:N
c 994      13.8      81.2      490      9      DA836585      DA836585
c 995      13.8      81.2      490      9      DB128459      DB128459
c 996      13.8      81.2      491      2      BI536042      390206 MA
c 997      13.8      81.2      491      7      BE979549      UI-M-RG2-
c 998      13.8      81.2      491      9      DA066275      DA066275
c 999      13.8      81.2      491      9      DA129611      DA129611
c1000     13.8      81.2      491      9      DA259777      DA259777

ALIGNMENTS

RESULT 1
BI063311      158 bp      mRNA      linear      EST 15-JUN-2001
LOCUS      UI3-UT0117-300301-538-E01 UT0117 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BI063311
ACCESSION      BI063311
VERSION      BI063311.1 GI:14470838
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE      1 (bases 1 to 158)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-UT0117-
300301-538-E01&t3=2001-03-30&t4=1)
Seq primer: puc 18 forward

FEATURES
source
1..307
/organism="Homo sapiens"
/db_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ; breast normal; Vector: puc18; Site 1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products

High quality sequence stop: 158.
Location/Qualifiers
1..158
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UT0117"
/note="Organ; uterus_tumor; Vector: puc18; Site 1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No.196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match      100.0%; Score 17; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No.5.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GCCACGCTTGGCGGAG 17
        |||||||
Db      100 GCCACGCTTGGCGGAG 116

RESULT 2
BE818833/c
LOCUS      CM2-BN0302-050700-256-c01 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BE818833
ACCESSION      BE818833
VERSION      BE818833.1 GI:10251067
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE      1 (bases 1 to 307)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2-BN0302-050
700-256-c01&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 307.

FEATURES
source
1..307
/organism="Homo sapiens"
/db_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ; breast normal; Vector: puc18; Site 1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
```

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 307;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
 |||||
 Db 242 GCCCAGCTTGGCCGAG 226

RESULT 3

BE818817/c
 LOCUS A112374
 DEFINITION CM2-BN0302-050700-256-b07 BN0302 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE818817
 VERSION BE818817.1 GI:10251051
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 341)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, P., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 341)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, P., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED

10737800

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=CM2-BN0302-050

700-256-b07&t3=2000-07-05&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 341.

FEATURES

source

1..341
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0302"

/note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 341;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
 |||||
 Db 253 GCCCAGCTTGGCCGAG 237

RESULT 4

A112374

LOCUS A112374

DEFINITION

zn68e10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone

IMAGE:563370 5', mRNA sequence.

ACCESSION A112374

VERSION A112374.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 367)

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

8889549

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LINDL; contact the

IMAGE Consortium (info@image.lindl.gov) for further information.

Insert Length: 1363 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 329.

Location/Qualifiers

1..367

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:4593181"

/db_xref="taxon:9606"

/clone="IMAGE:563370"

/sex="female"

/dev_stage="HeLa S3 cell line"

/clone_lib="SOLR (kanamycin resistant)"

/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. HeLa S3 epithelioid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR Vector. -5' adaptor sequence: 5' GAATTCGCGCAG 3' -3' adaptor sequence: 5' CTCGATTTTTTTTTTTTTTTT 3"

Query Match 100.0%; Score 17; DB 1; Length 367;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

|||||

Db 116 GCCCAGCTTGGCCGAG 132

RESULT 5

BE818835/c

LOCUS BE818835

DEFINITION CM2-BN0302-050700-256-c11 BN0302 Homo sapiens cDNA, mRNA sequence.

```

ACCESSION      BE818835
VERSION        BE818835.1  GI:10251069
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 384)
AUTHORS        Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
              O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
              10737800
              Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=CM2-BN0302-050
              700-256-clitk3=2000-07-05&t4=1)
              Seq primer: puc 18 forward
              High quality sequence start: 5
              High quality sequence stop: 384.
FEATURES       source
               1..384
                Location/Qualifiers
                1..384
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /dev_stage="Adult"
                /clone_lib="BN0302"
                /notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
                Site_2: SmaI; A mini-library was made by cloning products
                derived from ORESTES PCR (U.S. Letters Patent application
                No. 196,716 - Ludwig Institute for Cancer Research)
                profiles into the pUC 18 vector. Reverse transcription of
                tissue mRNA and cDNA amplification were performed under
                low stringency conditions."
ORIGIN
Query Match      100.0%; Score 17; DB 7; Length 384;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
   ||||||||||||||||
Db 257 GCCCAGCGTTGGCCGAG 241

RESULT 6
BE818828/c
LOCUS          BE818828          396 bp      mRNA      linear      EST 21-SEP-2000
DEFINITION    CM2-BN0302-050700-256-a05 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION     BE818828
VERSION       BE818828.1  GI:10251062
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 396)
AUTHORS        Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
              O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
              10737800
              Contact: Simpson A.J.G.

```

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimposon@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-CM2-BN0302-050>)
700-256-f02&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 404.

FEATURES

Location/Qualifiers
source
1..404
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

/note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

|||||
Db 244 GCCCAGCGTTGGCCGAG 228

RESULT 8

BE818805/c
LOCUS
DEFINITION CM2-BN0302-040700-253-b10 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE818805
VERSION
BE818805.1 GI:10251039
KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1 (bases 1 to 424)

Authors: Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED
10737800

COMMENT
Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimposon@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-CM2-BN0302-040>)

700-253-b10&t3=2000-07-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 424.

FEATURES

Location/Qualifiers
source
1..424
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 424;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

|||||
Db 250 GCCCAGCGTTGGCCGAG 234

RESULT 9

CD672930
LOCUS
DEFINITION CD672930 453 bp mRNA linear EST 24-JUN-2003
fg18a05.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
fg18a05.5', mRNA sequence.

ACCESSION
CD672930

VERSION
CD672930.1 GI:32174661

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1 (bases 1 to 453)

Wistow, G., Bernstein, S.L., Ray, S., Wyatt, M.K., Behal, A.,

Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.

Expressed sequence tag analysis of adult human iris for the NEtBank

Project: steroid-response factors and similarities with retinal

pigment epithelium

Mol. Vis. 8 (4), 185-195 (2002)

12107412

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 18 row: a column: 05

Seq primer: M13RPL reverse primer (ABI).

FEATURES

Location/Qualifiers
source
1..453
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="fg18a05"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDHL0B"
/clone_lib="Human Iris cDNA (Normalized): fg"

/note="Organ: Eye; Vector: pCMVSPORT6; A human iris
library (bx) was normalized by self-subtraction. One
portion of double stranded plasmid DNA representing the
library was linearized by NotI. This NotI digested library

was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (Cot 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 100.0%; Score 17; DB 5; Length 453;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCGAG 17
|||||
Db 56 GCCACGCTTGCGCGAG 72

RESULT 10

BE818871/c
LOCUS BE818871 458 bp mRNA linear EST 21-SEP-2000
DEFINITION CM2-BN0302-100700-259-f01 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE818871
VERSION BE818871.1 GI:10251105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

1 (bases 1 to 458)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed

TITLE

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800

JOURNAL
PUBMED

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-100700-259-f01&t3=2000-07-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 12

High quality sequence stop: 458.

FEATURES

source

1. 458
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="BN0302"
/note="Organ: breast normal; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 458;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCGAG 17
|||||
Db 264 GCCACGCTTGCGCGAG 248

RESULT 11

BE818859/c
LOCUS BE818859 466 bp mRNA linear EST 21-SEP-2000
DEFINITION CM2-BN0302-100700-259-c02 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE818859
VERSION BE818859.1 GI:10251093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

1 (bases 1 to 466)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed

TITLE

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800

JOURNAL
PUBMED

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-100700-259-c02&t3=2000-07-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 16

High quality sequence stop: 466.

FEATURES

source

1. 466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ: breast normal; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCGCGAG 17
|||||
Db 267 GCCACGCTTGCGCGAG 251

```
RESULT 12
BE818852/c
LOCUS      BE818852      467 bp      mRNA      linear      EST 21-SEP-2000
DEFINITION CM2-BN0302-100700-259-a05 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE818852
VERSION     BE818852.1      GI:10251086
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
            1 (bases 1 to 467)
REFERENCE   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
AUTHORS     Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brustein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM2-BN0302-100
            700-259-a05&t3=2000-07-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 467.
            High quality sequence stop: 467.
            Location/Qualifiers
            1..467
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="BN0302"
            /note="Organ: breast_normal; Vector: puc18; Site 1: SmaI;
            Site 2: SmaI; A mini-library was made by cloning products
            derived from ORFESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."

FEATURES             source
Query Match          100.0%; Score 17; DB 7; Length 467;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCCACGCTTGGCCGAG 17
    |||||
Db 246 GCCCACGCTTGGCCGAG 230

RESULT 13
AW390096
LOCUS      AW390096      528 bp      mRNA      linear      EST 04-FEB-2000
DEFINITION CM3-ST0179-051099-019-f07 ST0179 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW390096
VERSION     AW390096.1      GI:6894755
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
            1 (bases 1 to 528)
REFERENCE   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
AUTHORS     Mizushima-Sugano,J., Nakai,K. and Sugano,S.
            Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
            Genome Res. 14 (9), 1711-1718 (2004)
            15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Tel: 81-3-5449-5343
            Fax: 81-3-5449-5416
```


Email: yeuzuki@hgc.jp.
Location/Qualifiers
source
1. .571
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC03720"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 571;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCGGAG 17
|||||
Db 522 GCCCAGCGTTGGCGGAG 538

RESULT 15
CB152636 577 bp mRNA linear EST 29-JAN-2003
LOCUS K-EST0209868 L12JSHC0e1 Homo sapiens cDNA clone L12JSHC0e1-5-G07
DEFINITION 5', mRNA sequence.
ACCESSION CB152636
VERSION CB152636.1 GI:28137590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 577)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 5 row: G column: 07
High quality sequence stop: 577.
Location/Qualifiers
source
1. .577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L12JSHC0e1-5-G07"
/sex="M"
/cell_line="J-SHC"
/lab_host="Top10F"
/clone_lib="L12JSHC0e1"
/note="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.

The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dt)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the substracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

ORIGIN
Query Match 100.0%; Score 17; DB 4; Length 577;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCGGAG 17
|||||
Db 197 GCCCAGCGTTGGCGGAG 213

RESULT 16
BP267556 578 bp mRNA linear EST 16-SEP-2004
LOCUS BP267556 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
DEFINITION JTH07367, mRNA sequence.
ACCESSION BP267556
VERSION BP267556.1 GI:52182788
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 578)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.
Location/Qualifiers
source
1. .578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH07367"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/note="thyroid tumor"

ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCGGAG 17
|||||
Db 556 GCCCAGCGTTGGCGGAG 572

```
RESULT 17
BP265234
LOCUS
DEFINITION BP265234 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
ACCESSION JTH00322, mRNA sequence.
VERSION BP265234
KEYWORDS BP265234.1 GI:52180465
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
1 (bases 1 to 580)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
SEQUENCE COMPARISON OF HUMAN AND MOUSE GENES REVEALS A HOMOLOGOUS
BLOCK STRUCTURE IN THE PROMOTER REGIONS
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yuzuki@hgc.jp.
FEATURES
source
location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH00322"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"
ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 327 GCCCAGCGTTGGCCGAG 343

RESULT 18
BP267627
LOCUS
DEFINITION BP267627 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
ACCESSION JTH07584, mRNA sequence.
VERSION BP267627
KEYWORDS BP267627.1 GI:52182859
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
1 (bases 1 to 580)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
SEQUENCE COMPARISON OF HUMAN AND MOUSE GENES REVEALS A HOMOLOGOUS
BLOCK STRUCTURE IN THE PROMOTER REGIONS
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
```

```
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yuzuki@hgc.jp.
FEATURES
source
location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH07584"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"
ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 539 GCCCAGCGTTGGCCGAG 555

RESULT 19
BP268151
LOCUS
DEFINITION BP268151 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
ACCESSION JTH09144, mRNA sequence.
VERSION BP268151
KEYWORDS BP268151.1 GI:52183383
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
1 (bases 1 to 580)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
SEQUENCE COMPARISON OF HUMAN AND MOUSE GENES REVEALS A HOMOLOGOUS
BLOCK STRUCTURE IN THE PROMOTER REGIONS
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yuzuki@hgc.jp.
FEATURES
source
location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH09144"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"
ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 538 GCCCAGCGTTGGCCGAG 554

RESULT 20
```

BP256144
LOCUS BP256144 582 bp mRNA linear EST 16-SEP-2004
DEFINITION cDNA clone HRC10671, kidney epithelial cell Homo sapiens
ACCESSION BP256144
VERSION BP256144.1 GI:52171374
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yusuzuki@hgc.jp.

FEATURES
source
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC10671"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 415 GCCCAGCGTTGGCCGAG 431

RESULT 21
BP267918
LOCUS BP267918 582 bp mRNA linear EST 16-SEP-2004
DEFINITION BP267918 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
JTH08419, mRNA sequence.
ACCESSION BP267918
VERSION BP267918.1 GI:52183150
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yusuzuki@hgc.jp.

FEATURES
source
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH08419"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/note="thyroid tumor"

ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 556 GCCCAGCGTTGGCCGAG 572

RESULT 22
BP363758
LOCUS BP363758 582 bp mRNA linear EST 17-SEP-2004
DEFINITION BP363758 Sugano cDNA library, fetal lung fibroblast TIG Homo
sapiens cDNA clone TIR02383, mRNA sequence.
ACCESSION BP363758
VERSION BP363758.1 GI:52293963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yusuzuki@hgc.jp.

FEATURES
source
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TIR02383"
/tissue_type="lung"
/cell_type="fibroblast"
/cell_line="TIG"
/dev_stage="fetal"
/clone_lib="Sugano cDNA library, fetal lung fibroblast
TIG"

ORIGIN
Query Match 100.0%; Score 17; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 380 GCCCAGCGTTGGCCGAG 396

RESULT 23
BE006062

LOCUS BE006062 583 bp mRNA linear EST 05-JUN-2000
 DEFINITION RC2-BN0123-170300-011-c07 BN0123 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE006062
 VERSION BE006062.1 GI:8266295
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC2-BN0123-170300-011-c07&t3=2000-03-17&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 24
 High quality sequence stop: 583.
 Location/Qualifiers
 source
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0123"
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SnaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Query Match 100.0%; Score 17; DB 7; Length 583;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
 |||||
 Db 552 GCCCAGCTTGCCGAG 568

RESULT 24
 BP256147
 LOCUS BP256147 584 bp mRNA linear EST 16-SEP-2004
 DEFINITION cDNA clone HRC10676, mRNA sequence.
 ACCESSION BP256147
 VERSION BP256147.1 GI:52171377
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.
 1 (bases 1 to 584)
 Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 15342556
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: yezukui@bgc.jp.
 Location/Qualifiers
 source
 1..584
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HRC10676"
 /tissue_type="kidney"
 /cell_type="epithelial cell"
 /clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
 Query Match 100.0%; Score 17; DB 3; Length 584;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
 |||||
 Db 416 GCCCAGCTTGCCGAG 432

RESULT 25
 CB130950
 LOCUS CB130950 621 bp mRNA linear EST 29-JAN-2003
 DEFINITION K-EST0180946 L12JSHCO Homo sapiens cDNA clone L12JSHCO-5-D03 5', mRNA sequence.
 ACCESSION CB130950
 VERSION CB130950.1 GI:28095412
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 621)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 5 row: D column: 03
 High quality sequence stop: 621.
 Location/Qualifiers
 source
 1..621
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L12JSHCO-5-D03"
 /sex="M"
 /cell_line="J-SHC"
 /lab_host="Top10F"
 /clone_lib="L12JSHCO"

/note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 100.0%; Score 17; DB 4; Length 621;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 314 GCCCAGCGTTGGCCGAG 330

RESULT 26
BE747107
LOCUS BE747107 630 bp mRNA linear EST 15-SEP-2000
DEFINITION 601580680F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3929706 5', mRNA sequence.

ACCESSION BE747107
VERSION BE747107.1 GI:10161099
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 630)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICMT62 row: 0 column: 19
High quality sequence stop: 628.

FEATURES

source

1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3929706"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"

/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 100.0%; Score 17; DB 7; Length 630;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 66 GCCCAGCGTTGGCCGAG 82

RESULT 27

BM843051

LOCUS

DEFINITION

K-EST0120649 S12SNU216 Homo sapiens cDNA clone S12SNU216-77-H12 5', mRNA sequence.

ACCESSION

BM843051

VERSION

BM843051.1 GI:19199460

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 637)

AUTHORS

Oh, J.H., Yang, J.O., Hahn, Y., Kim, M.R., Byun, S.S., Jeon, Y.J., Kim, J.M., Song, K.S., Noh, S.M., Kim, S., Yoo, H.S., Kim, Y.S. and Kim, N.S.

TITLE

Transcriptome analysis of human gastric cancer

JOURNAL

Mamm. Genome 16 (12), 942-954 (2005)

PUBMED

16341674

COMMENT

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 77 row: H column: 12

High quality sequence stop: 637.

Location/Qualifiers

1. .637

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S12SNU216-77-H12"

/sex="F"

/tissue_type="Lymph node"

/cell_type="Epithelial"

/cell_line="SNU-216"

/lab_host="Top10F"

/clone_lib="S12SNU216"

/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 100.0%; Score 17; DB 3; Length 637;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17

```

Db      59  GCCCAGCGTTGGCCGAG 75
|||||
REFERENCE
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-remail.nih.gov
         Tissue Procurement: ATCC
         cDNA Library Preparation: Ling Hong/Rubin Laboratory
         DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Incyte Genomics, Inc.
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM158 row: 0 column: 01
         High quality sequence stop: 638.
FEATURES
source   Location/Qualifiers
1. .699
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300248"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
       Site 2: EcoRI; cDNA made by oligo-dT priming.
       Directionally cloned into EcoRI/XhoI sites using the
       following 5' adaptor: GCCAGCGAG(G). Size-selected >50bp
       for average insert size 1.8kb. Library constructed by Ling
       Hong in the laboratory of Gerald M. Rubin (University of
       California, Berkeley) using ZAP-cDNA synthesis kit
       (Stratagene) and Superscript II RT (Life Technologies).
       Note: this is a NIH_MGC Library."
ORIGIN
Query Match 100.0%; Score 17; DB 7; Length 699;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  GCCCAGCGTTGGCCGAG 17
|||||

Db 321 GCCCAGCGTTGGCCGAG 337
|||||

RESULT 29
BG251131
LOCUS    602364932F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473439 5',
DEFINITION mRNA sequence.
ACCESSION BG251131
VERSION    BG251131.1 GI:12760947
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.

REFERENCE
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-remail.nih.gov
         Tissue Procurement: ATCC
         cDNA Library Preparation: Ling Hong/Rubin Laboratory
         DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Incyte Genomics, Inc.
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM158 row: 0 column: 01
         High quality sequence stop: 638.
FEATURES
source   Location/Qualifiers
1. .699
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300248"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
       Site 2: EcoRI; cDNA made by oligo-dT priming.
       Directionally cloned into EcoRI/XhoI sites using the
       following 5' adaptor: GCCAGCGAG(G). Size-selected >50bp
       for average insert size 1.8kb. Library constructed by Ling
       Hong in the laboratory of Gerald M. Rubin (University of
       California, Berkeley) using ZAP-cDNA synthesis kit
       (Stratagene) and Superscript II RT (Life Technologies).
       Note: this is a NIH_MGC Library."
ORIGIN
Query Match 100.0%; Score 17; DB 7; Length 699;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  GCCCAGCGTTGGCCGAG 17
|||||

Db 321 GCCCAGCGTTGGCCGAG 337
|||||

RESULT 29
BG251131
LOCUS    602364932F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473439 5',
DEFINITION mRNA sequence.
ACCESSION BG251131
VERSION    BG251131.1 GI:12760947
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.

REFERENCE
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-remail.nih.gov
         Tissue Procurement: ATCC
         cDNA Library Preparation: Ruben Laboratory
         DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Agencourt Bioscience Corporation
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM2570 row: 1 column: 09
         High quality sequence stop: 579.
FEATURES
source   Location/Qualifiers
1. .851
/mol_type="mRNA"
/organism="Homo sapiens"

```

```

Hominidae; Homo.
1 (bases 1 to 702)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-remail.nih.gov
         Tissue Procurement: ATCC
         cDNA Library Preparation: Life Technologies, Inc.
         DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Incyte Genomics, Inc.
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM10295 row: 0 column: 08
         High quality sequence stop: 688.
FEATURES
source   Location/Qualifiers
1. .702
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4473439"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
       Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
       Average insert size 1.7 kb. Library enriched for
       full-length clones and constructed by Life Technologies.
       Note: this is a NIH_MGC Library."
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 702;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  GCCCAGCGTTGGCCGAG 17
|||||

Db 464 GCCCAGCGTTGGCCGAG 480
|||||

RESULT 30
BQ887352
LOCUS    BQ887352
DEFINITION AGENCOURT 8678064 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381080
5', mRNA sequence.
ACCESSION BQ887352
VERSION    BQ887352.1 GI:22279366
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.

REFERENCE
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
         Email: cgapbs-remail.nih.gov
         Tissue Procurement: DCTD/DTF
         cDNA Library Preparation: Ruben Laboratory
         DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Agencourt Bioscience Corporation
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM2570 row: 1 column: 09
         High quality sequence stop: 579.
FEATURES
source   Location/Qualifiers
1. .851
/mol_type="mRNA"
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
 /clone="IMAGE:6381080"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCCAGG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 17; DB 3; Length 851;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAG 17
 |||||
 Db 182 GCCACGCTTGGCCGAG 198

RESULT 31
 CA454746
 LOCUS
 DEFINITION AGENCOURT 10763191 MAPcL Homo sapiens cDNA clone IMAGE:6721245 5',
 mRNA sequence.

ACCESSION CA454746
 VERSION GI:24904781
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.

REFERENCE 1 (bases 1 to 865)
 NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov
 Tissue Procurement: Kristi A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM14282 row: a column: 21

High quality sequence stop: 686.

Location/Qualifiers

FEATURES

1..865

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6721245"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,

HTERT-HME1, LNCaP"

/lab_host="EMDH10B"

/clone_lib="MAPcL"

/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
 Subtracted with brain, liver, lung, kidney and muscle.

Directionally cloned. Priming method: oligo-dT. Average
 insert size: 1800 bp. Library amplification: 26,000 fold.

Kristi A. Eglund, James J. Vincent, Robert Strausberg,
 Bungkok Lee & Ira Pastan. Discovery of new breast

cancer genes encoding membrane and secreted proteins.

Manuscript submitted."

ORIGIN

Query Match 100.0%; Score 17; DB 4; Length 865;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAG 17
 |||||
 Db 611 GCCACGCTTGGCCGAG 627

RESULT 32
 BM011054
 LOCUS
 DEFINITION BM011054
 603634744F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5421755 5',
 mRNA sequence.

ACCESSION BM011054

VERSION GI:16525408

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.

REFERENCE 1 (bases 1 to 877)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LCM1878 row: h column: 12

High quality sequence stop: 746.

Location/Qualifiers

FEATURES

1..877

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5421755"

/tissue_type="neuroblastoma, cell line"

/lab_host="NIH_MGC_47"

/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCCAGG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 17; DB 2; Length 877;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAG 17
 |||||
 Db 184 GCCACGCTTGGCCGAG 200

RESULT 33
 BQ958341
 LOCUS
 DEFINITION AGENCOURT 10037048 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6480481
 5', mRNA sequence.

ACCESSION BQ958341

VERSION BQ958341.1 GI:22373819

KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 899)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2660 row: f column: 02
 High quality sequence stop: 636.

FEATURES
source
 1..899
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6480481"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 100.0%; Score 17; DB 3; Length 899;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCCGAG 17
 |||||

Db 546 GCCCACGCTTGGCCGAG 562
 |||||

RESULT 34
LOCUS BG831564
DEFINITION BG831564.1 GI:14179151
 mRNA sequence.
ACCESSION BG831564
VERSION BG831564.1 GI:14179151
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 901)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: ATCC
 Email: cgapbs-r@mail.nih.gov
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: L1CM1813 row: e column: 06
 High quality sequence stop: 663.

FEATURES
source
 1..901
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4907885"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_42"
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 100.0%; Score 17; DB 2; Length 901;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCCGAG 17
 |||||

Db 322 GCCCACGCTTGGCCGAG 338
 |||||

RESULT 35
LOCUS CF780868
DEFINITION CF780868 916 bp mRNA linear EST 20-OCT-2003
 AGENCOURT 15939245 NIH MGC 219 Homo sapiens cDNA clone
 IMAGE:30523569 5', mRNA sequence.

ACCESSION CF780868
VERSION CF780868.1 GI:37740645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 916)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM601 row: l column: 10
 High quality sequence stop: 651.

FEATURES
source
 1..916
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30523569"
 /tissue_type="Pooled Chondrosarcoma Tumor cells"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_219"
 /note="Vector: pYX-Asc; Site 1: EcoRI; Site 2: NotI;

Library is oligo-dT primed and directionally cloned. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 0.5-1kb. Adaptors 5' (AATTCGGCAGG)3' and 5' (CTCGTGGCG)3'. 3' linker sequence - GCGCGCGTCGAGGCC T18. Sequencing primers 3' end: T3 promoter primer 5' (ATTAACCTCACTTAAGGA)3'. 5' End: T7 promoter primer 5' (TAATACGACTCATATAGG)3'. Library was constructed in the laboratory of M. Bento Soares. Average insert size 2-3kb. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 17; DB 5; Length 916;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 667 GCCCAGCGTTGGCCGAG 683

RESULT 36

BX372160/c

LOCUS

DEFINITION BX372160 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK011YF17 3-PRIME, mRNA sequence.

ACCESSION

BX372160

VERSION

BX372160.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 918)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segret@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

6601.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0BA1040ZB07_CS03797_14c=6601.f

Location/Qualifiers

1. .918

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="CS0DK011YF17"

/cell_type="HELA CELLS COT 25-NORMALIZED"

/clone_lib="HELA"

/note="First strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 17; DB 4; Length 918;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 GCCCAGCGTTGGCCGAG 17

Db

335 GCCCAGCGTTGGCCGAG 319

RESULT 37

BU543952

LOCUS

DEFINITION

BU543952

5', mRNA sequence.

ACCESSION

BU543952

VERSION

BU543952.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 923)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM2774 row: d column: 08

High quality sequence stop: 641.

Location/Qualifiers

1. .923

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6576152"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 17; DB 3; Length 923;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 GCCCAGCGTTGGCCGAG 17

Db

546 GCCCAGCGTTGGCCGAG 562

RESULT 38

BX370558

LOCUS

DEFINITION

BX370558

Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens

cDNA clone CS0DK011YF17 5-PRIME, mRNA sequence.

ACCESSION

BX370558

VERSION

BX370558.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 962)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6601.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAG0492C03_CS04627_lkc=6601.f

FEATURES
source

Location/Qualifiers
1..962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YF17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 17; DB 4; Length 962;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTGGCCGAG 17
|||||
Db 248 GCCCAGCTGGCCGAG 264

RESULT 39
BQ067433
LOCUS
DEFINITION BQ067433 1023 bp mRNA linear EST 02-APR-2002
AGENCOURT 6758944 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754944
5', mRNA sequence.
ACCESSION BQ067433
VERSION BQ067433.1 GI:19896479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1023)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM12793 row: c column: 09
High quality sequence stop: 649.
Location/Qualifiers
1..1023

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM12793 row: c column: 09
High quality sequence stop: 649.
Location/Qualifiers
1..1023

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5754944"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 17; DB 3; Length 1023;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTGGCCGAG 17
|||||
Db 403 GCCCAGCTGGCCGAG 419

RESULT 40
BI193620
LOCUS
DEFINITION BI193620 1125 bp mRNA linear EST 10-JUL-2001
602946519F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5089900 5',
mRNA sequence.
ACCESSION BI193620
VERSION BI193620.1 GI:14648640
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 1125)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1851 row: e column: 05
High quality sequence stop: 529.
Location/Qualifiers
1..1125

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5089900"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="NIH_MGC_42"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

ORIGIN
Query Match      100.0%; Score 17; DB 2; Length 1125;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCGCGAG 17
    |||||
Db 294 GCCACGCTTGCGCGAG 310

RESULT 41
CR597125
LOCUS
DEFINITION full-length cDNA clone CS0DK011YF17 of HeLa cells Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR597125
VERSION CR597125.1 GI:50477932
KEYWORDS HTC; CNSUT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1830)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1830)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1..1830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YF17"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match      100.0%; Score 17; DB 6; Length 1830;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCGCGAG 17
    |||||
Db 1118 GCCACGCTTGCGCGAG 1134

RESULT 42
AA041289/c
LOCUS
DEFINITION zf08d03.e1 Soares fetal heart NBHL19W Homo sapiens cDNA clone
IMAGE376325 3' similar to gb:X02162 AF010PROTEIN A-I PRECURSOR
(HUMAN); mRNA sequence.
ACCESSION AA041289
VERSION AA041289.1 GI:1517506
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 322)
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 288.
FEATURES
source
Location/Qualifiers
1..322
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1284581"
/db_xref="taxon:9606"
/clone="IMAGE:376325"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NBHL19W"
/noted="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCATCTGAAGTGGGAGCGCGCATCTTTTTTTTTTTTTTT 3'], [5',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."

ORIGIN
Query Match      90.6%; Score 15.4; DB 1; Length 322;
Best Local Similarity 94.1%; Pred. No. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACGCTTGCGCGAG 17
    |||||
Db 270 GCCACGCTTGCGCGAG 254

RESULT 43
BY153767/c
LOCUS
DEFINITION BY153767 RIKEN full-length enriched, adult pancreas islet cells Mus
musculus cDNA clone C820014J12 5', mRNA sequence.
ACCESSION BY153767
VERSION BY153767.1 GI:26290313
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 363)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

```

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lethard, B., Lyons, P.A., Maglott, D.R., Maltate, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL
PUBMED**
Nature 420, 563-573 (2002)
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1517-1530 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Cells were provided by Hiroo Iwata (Department of Reproductive Materials Field of Tissue Engineering Institute for Frontier Medical Sciences, Sakyo-ku, Kyoto, 606-8507, Japan) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
FEATURES
source
Location/Qualifiers
1. .363
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="C820014J12"
/tissue_type="pancreas"
/cell_type="islet cells"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult pancreas

islet cells"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGCGCGCCCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTTAATAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cells were provided by Tomohiro Kono, Department of Animal Science, Tokyo University of Agriculture, Atsugi City, Kanagawa Prefecture, Japan, whose assistance we gratefully acknowledge."

ORIGIN

Query Match 90.6%; Score 15.4; DB 4; Length 363;
Best Local Similarity 94.1%; Pred. No. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACCCTTGGCCGAG 17
|||||
Db 206 GCCCGCGCTTGGCCGAG 190
|||||

RESULT 44
BY774337/c
LOCUS BY774337 365 bp mRNA linear EST 23-MAR-2004
DEFINITION BY774337 RIKEN full-length enriched, 17.5 days embryo whole body
Mus musculus cDNA clone L930083C03 5', mRNA sequence.
BY774337
ACCESSION BY774337.1 GI:39700975
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 365)
Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aidinis, V., Nakagawara, A., Held, W.A., Iwata, H., Kono, T., Nakachi, H., Lyons, P., Wells, C., Hume, D.A., Fagioli, M., Hensch, T.K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia
Genome Res. 13 (6B), 1273-1289 (2003)
12819125
JOURNAL
PUBMED
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

```

FEATURES
  source
    Location/Qualifiers
      1. 365
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="L930083C03"
        /tissue_type="whole body"
        /dev_stage="17.5 days embryo"
        /clone_lib="RIKEN full-length enriched, 17.5 days embryo
        whole body"

ORIGIN
  Query Match      90.6%; Score 15.4; DB 4; Length 365;
  Best Local Similarity 94.1%; Pred. No. 3.6e+03;
  Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

  QY 1 GCCACGCTTGCGCGAG 17
      ||||| ||||| |||||
  Db 164 GCCCGCGCTTGCGCGAG 148

RESULT 45
BE818831/c
LOCUS
DEFINITION
  B818831 365 bp mRNA linear EST 21-SEP-2000
ACCESSION
  B818831.1 GI:10251065
KEYWORDS
  EST.
ORGANISM
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominoidea; Homo.
REFERENCE
  1 (bases 1 to 365)
  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
  Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
  Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
  Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
  O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
  Simpson,A.J.
  Shotgun sequencing of the human transcriptome with ORF expressed
  sequence tags
  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
COMMENT
  Contact: Simpson A.J.G.
  Laboratory of Cancer Genetics
  Ludwig Institute for Cancer Research
  Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
  Brazil
  Tel: +55-11-2704922
  Fax: +55-11-2707001
  Email: asimpson@ludwig.org.br
  This sequence was derived from the FAPESP/LICR Human Cancer Genome
  Project. This entry can be seen in the following URL
  (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-EN0302-050
  700-256-a12&t3=2000-07-05&t4=1)
  Seq primer: puc 18 forward
  High quality sequence start: 14
  High quality sequence stop: 365.
  Location/Qualifiers
    1. 365
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
      /clone_lib="EN0302"
      /notes="Organ: breast normal; Vector: puc18; Site:1: Smal;
      Site 2: SmaI; A mini-library was made by cloning products
      derived from ORESTES PCR (U.S. Letters Patent application
      No. 196, 716 - Ludwig Institute for Cancer Research)
      profiles into the pUC 18 vector. Reverse transcription of
      tissue mRNA and cDNA amplification were performed under

FEATURES
  source
    Location/Qualifiers
      1. 365
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="L930083C03"
        /tissue_type="whole body"
        /dev_stage="17.5 days embryo"
        /clone_lib="RIKEN full-length enriched, 17.5 days embryo
        whole body"

ORIGIN
  Query Match      90.6%; Score 15.4; DB 7; Length 365;
  Best Local Similarity 94.1%; Pred. No. 3.6e+03;
  Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

  QY 1 GCCACGCTTGCGCGAG 17
      ||||| ||||| |||||
  Db 267 GCCACGCTTGCGCGAG 251

RESULT 46
BY674396
LOCUS
DEFINITION
  BY674396 RIKEN full-length enriched, 14.5 days embryo df/df
  Rathke's pouches Mus musculus cDNA clone K820040J08 3', mRNA
  sequence.
  BY674396
  BY674396.1 GI:27058918
  EST.
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 367)
  Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
  Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaoka,I.,
  Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
  Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
  Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
  Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
  Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
  Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
  Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
  Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
  Kawai,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Kongagaya,A.,
  Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
  Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
  Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
  Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
  Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
  Sandelin,A., Schneider,C., Sempie,C.A., Setou,M., Shimada,K.,
  Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
  Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
  Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
  Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
  Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
  Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
  Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
  Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
  Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
  Rogers,J., Birney,E. and Hayashizaki,Y.
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
  Nature 420, 563-573 (2002)
12466851
COMMENT
  Contact: Yoshihide Hayashizaki
  Laboratory for Genome Exploration Research Group, RIKEN Genomic
  Sciences Center(GSC), Yokohama Institute
  The Institute of Physical and Chemical Research (RIKEN)
  1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  Tel: 81-45-503-9222
  Fax: 81-45-503-9216
  Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
  Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
  Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
  Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
  Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
  Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
  Hayashizaki,Y. Direct Submission
  Computational Analysis of Full-Length Mouse cDNAs Compared with

```


Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301 MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

```
1..367
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K82004J08"
/tissue_type="Rathke's pouches"
/dev_stage="14.5 days embryo df/df"
/clone_lib="RIKEN full-length enriched, 14.5 days embryo
df/df Rathke's pouches"
```

ORIGIN

Query Match 90.6%; Score 15.4; DB 4; Length 367;
 Best Local Similarity 94.1%; Pred. No. 3.6e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCCGAG 17

Db 259 GCACACGCTTGCCCGAG 275

RESULT 47

BY674625

LOCUS

DEFINITION BY674625 RIKEN full-length enriched, 14.5 days embryo df/df
 Rathke's pouches Mus musculus cDNA clone K82004J17 3', mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 383)

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yananaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L. B., Cousins, S., Dalla, E., Draganic, T. A.,
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
 Maitais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W. J., Perlea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yangisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851

JOURNAL

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-3216

Email: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Wataniki, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301 MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

```
1..383
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K82004J17"
/tissue_type="Rathke's pouches"
/dev_stage="14.5 days embryo df/df"
/clone_lib="RIKEN full-length enriched, 14.5 days embryo
df/df Rathke's pouches"
```

ORIGIN

Query Match 90.6%; Score 15.4; DB 4; Length 383;
 Best Local Similarity 94.1%; Pred. No. 3.6e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCCGAG 17

Db 275 GCACACGCTTGCCCGAG 291

RESULT 48


```
AW206921/c
LOCUS
DEFINITION
  UI-H-B11-afg-d-01-0-UI.s1 NCI CGAP_Sub3 Homo sapiens cDNA clone
  IMAGE:2722968 3', mRNA sequence.
ACCESSION
  AW206921
VERSION
  AW206921.1 GI:6506417
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
REFERENCE
  1 (bases 1 to 410)
  NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapsb@mail.nih.gov
  The sequence contained an oligo-dT track that was present in the
  oligonucleotide that was used to prime the synthesis of first
  strand cDNA and therefore this may represent a bonafide poly A
  tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
  NCI-CCGAP clone distribution information can be found through the
  I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Seq primer: M13 Forward
  POLYA=Yes.
FEATURES
  source
    Location/Qualifiers
      1..410
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clones="IMAGE:2722968"
        /lab_hosts="DH10B (Life Technologies)"
        /clone_lib="NCI_CGAP_Sub3"
        /notes="vector: pVT73D-FacI; Site_1: Not I; Site_2: Eco RI;
        The NCI_CGAP_Sub3 library is a subtracted library derived
        from the NCI_CGAP_Sub1 library, which is a subtracted
        library derived from BI. BI constitutes a mixture of 21
        normalized or subtracted NCI_CGAP libraries:
        NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
        NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
        NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
        NCI_CGAP_Br23, NCI_CGAP_Co8, NCI_CGAP_C111, NCI_CGAP_Lei2,
        NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
        NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
        NCI_CGAP_Brn25. These 21 libraries were pooled and a
        single-stranded DNA preparation of the resulting mixture
        was used as a tracer in a subtractive hybridization with
        a driver whose composition is detailed below:
        NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
        3798-3803 (IMAGE CloneIDs 132376-132391),
        1456008-1456775, 1500552-1502855; NCI_CGAP_Kid5 pool 1
        LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs
        1323912-1325831, 1471368-1472903, 1492104-1493255);
        NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE
        CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4
        pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
        CloneIDs 1257096-1258631, 1469064-1470983,
        1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459,
        2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
        1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
        LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,
        1144584-1145351). Subtraction was performed as previously
        described [Bonaldi, Lennon & Soares (1996): Normalization
        and Subtraction: Two Approaches to Facilitate Gene
        Discovery. Genome Research 6, 791-806.
        TAG_TISSUE=colon
        TAG_LIB=NCI_CGAP_Co10
        TAG_SEQ=AAAGG"
ORIGIN
```

```
Query Match 90.6%; Score 15.4; DB 7; Length 410;
Best Local Similarity 94.1%; Pred. No. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCCGAG 17
   |||||
Db 359 GCCCAGCGCTTGGCCGG 343

RESULT 49
BE818807/c
LOCUS
DEFINITION
  CM2-BN0302-040700-253-d02 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION
  BE818807
VERSION
  BE818807.1 GI:10251041
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
REFERENCE
  1 (bases 1 to 449)
  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
  Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
  Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
  Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V.,
  O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
  Simpson,A.J.
  Shotgun sequencing of the human transcriptome with ORF expressed
  sequence tags
  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  10737800
  Contact: Simpson A.J.G.
  Laboratory of Cancer Genetics
  Ludwig Institute for Cancer Research
  Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
  Brazil
  Tel: +55-11-2704922
  Fax: +55-11-2707001
  Email: asimpson@ludwig.org.br
  This sequence was derived from the FAPESP/LICR Human Cancer Genome
  Project. This entry can be seen in the following URL
  (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-cm2-bn0302-040)
  700-253-d02&t3=2000-07-04&t4=1)
  Seq primer: puc 18 forward
  High quality sequence start: 24
  High quality sequence stop: 306.
FEATURES
  source
    Location/Qualifiers
      1..449
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /dev_stage="Adult"
        /clone_lib="BN0302"
        /notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
        Site_2: SmaI; A mini-library was made by cloning products
        derived from ORESTES PCR (U.S. Letters Patent application
        No. 136,716 - Ludwig Institute for Cancer Research)
        profiles into the pUC 18 vector. Reverse transcription of
        tissue mRNA and cDNA amplification were performed under
        low stringency conditions."
ORIGIN

Query Match 90.6%; Score 15.4; DB 7; Length 449;
Best Local Similarity 94.1%; Pred. No. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCCGAG 17
   |||||
Db 240 GCCCAGCGCTTGGCCGG 224

RESULT 50
```

DU607602
 LOCUS OO_Ba0106010.1 r OO_Ba Oryza officinalis genomic clone linear DNA 451 bp GSS 12-OCT-2005
 DEFINITION OO_Ba0106010.3, genomic survey sequence.
 ACCESSION DU607602
 VERSION DU607602.1 GI:77571132
 KEYWORDS GSS.
 SOURCE
 ORGANISM Oryza officinalis
 Oryza officinalis
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoideae; Oryzoideae; Oryza.
 REFERENCE 1 (bases 1 to 451)
 AUTHORS Kim, H., Collura, K., Wisotski, M., Byrne, M., Stum, D., Smart, D.,
 Rao, K., Luo, M., Jetty, R., Kudrna, D., Muller, C., Soderlund, C. and
 Wing, R.
 TITLE OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
 JOURNAL
 COMMENT Unpublished (2005)
 Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 PCR Primers
 FORWARD: TAA TAC GAC TCA CTA TAG GG
 BACKWARD: CAC TCA TTA GGC ACC CCA
 Plate: 0106 row: O column: 10
 Seq primer: CAC TCA TTA GGC ACC CCA
 Class: BAC ends.
 FEATURES
 source
 Location/Qualifiers
 1. 451
 /organism="Oryza officinalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:4535"
 /clone="OO_Ba0106010"
 /tissue_type="leaves"
 /lab_host="DH10B"
 /clone_lib="OO_Ba"
 /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
 ORIGIN
 Query Match 90.6%; Score 15.4; DB 14; Length 451;
 Best Local Similarity 94.1%; Pred. No. 3.6e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCCCAGCGTTGGCCGAG 17
 |||||
 Db 290 GCCCAGCGTTGGCCGTG 306
 RESULT 51
 BQ965719
 LOCUS BQ965719 494 bp mRNA linear EST 21-AUG-2002
 DEFINITION Oryza officinalis sunflower RHA801 Helianthus annuus cDNA
 clone QHB22M12, mRNA sequence.
 ACCESSION BQ965719
 VERSION BQ965719.1 GI:22382824
 KEYWORDS EST.
 SOURCE Helianthus annuus (common sunflower)
 ORGANISM Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 REFERENCE 1 (bases 1 to 494)
 AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Riesberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
 Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
 Lai, Z., Church, S., Jackson, L. and Bradford, K.
 TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/

Unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmunsden Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig980, see http://cgdb.ucdavis.edu/
 for details.
 Plate: QHB22 row: M column: 12.
 FEATURES
 source
 Location/Qualifiers
 1. 494
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA801"
 /db_xref="taxon:4232"
 /clone="QHB22M12"
 /lab_host="E.Coli"
 /clone_lib="QH ABCDI sunflower RHA801"
 /note="Vector: pBRCNDSFIAB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgdb.ucdavis.edu/
 TAG_SEQ=Not found"
 ORIGIN
 Query Match 90.6%; Score 15.4; DB 3; Length 494;
 Best Local Similarity 94.1%; Pred. No. 3.6e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCCCAGCGTTGGCCGAG 17
 |||||
 Db 465 GCCCAGCGTTGGCCGAG 481
 RESULT 52
 DA945234
 LOCUS DA945234 566 bp mRNA linear EST 15-NOV-2005
 DEFINITION SPLEN2 Homo sapiens cDNA clone SPLEN2020625 5', mRNA
 sequence.
 ACCESSION DA945234
 VERSION DA945234.1 GI:82439583
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 566)
 AUTHORS Kimura, K., Wakanatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 TITLE Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative
 Promoters of Human Genes
 JOURNAL Genome Res. 16 (1), 55-65 (2006)
 PUBMED 16344560
 COMMENT Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB. The reason of remaining 5'-end 'N': This sequence was generated from 5'-end region of cDNA insert including about 20 base of the sequence before 'N'. This cDNA was obtained from oligo-capped cDNA library by which method 5'-end of mRNA was efficiently cloned. Therefore remaining 5'-end 'N' is important for a detection of transcription start site of this mRNA.

FEATURES

source
1. .566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SPLEN2020625"
/tissue_type="spleen"
/clone_lib="SPLEN2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 90.6%; Score 15.4; DB 9; Length 566;
Best Local Similarity 94.1%; Pred. NO. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGGAG 17

Db 482 GCCACGCTTGGCGGAG 498

RESULT 53

DA932540 DA932540 SPLEN2 Homo sapiens cDNA clone SPLEN2002176 5', mRNA
LOCUS
DEFINITION

sequence.

DA932540 DA932540.1 GI:82096064

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Murakawa, K., Iehida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560

JOURNAL

PUBMED

COMMENT

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Karusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES

source
1. .570
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="SPLEN2002176"
/tissue_type="spleen"
/clone_lib="SPLEN2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 90.6%; Score 15.4; DB 9; Length 570;
Best Local Similarity 94.1%; Pred. NO. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACGCTTGGCGGAG 17

Db 459 GCCACGCTTGGCGGAG 475

RESULT 54

CK783403/c

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CK783403 577 bp mRNA linear EST 19-OCT-2004
UI-D-GC1-aau-g-20-0-UI.s1 UI-D-GC1 Alexandrium tamarense cDNA clone
UI-D-GC1-aau-g-20-0-UI 3', mRNA sequence.

CK783403

CK783403.1 GI:42749081

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

1 (bases 1 to 577)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
8889548

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu
Tissue Procurement: Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP)
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/dinoflagellate.html
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers

1. .577

/organism="Alexandrium tamarense"

/mol_type="mRNA"

/strain="CCMP 1598"

/db_xref="taxon:2926"

/clone="UI-D-GC1-aau-g-20-0-UI"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-D-GC1"

/note="Vector: pT7T3-Pac (Pharmacia) with a modified

polylinker; Site 1: EcoR I; Site 2: Not I; UI-D-GC1 is a

normalized library derived from UI-D-GC0. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pT7T3-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

TACTCTCGAGA. Tissue was obtained from the

Provasoli-Guillard National Center for Culture of Marine

Phytoplankton (CMP).
TAG TISSUE=Alexandrium tanarense
TAG_LIB=UI-D-GC1
TAG_SEQ=TACTCGAGA"

ORIGIN

Query Match 90.6%; Score 15.4; DB 5; Length 577;
Best Local Similarity 94.1%; Pred. No. 3.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
|||
Db 74 GCTCACGCTTGCCGAG 58
|||

RESULT 55

A2259233/c

LOCUS
DEFINITION RPCI-23-136M7.TV RPCI-23 Mus musculus genomic clone RPCI-23-136M7,
genomic survey sequence.

ACCESSION A2259233

VERSION A2259233.1 GI:9465509

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 582)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Teegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.

TITLE

JOURNAL

COMMENT

Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)

Other GSSs: RPCI-23-136M7.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BAC PAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)or from Resea ch Genetics (info@resgen.com). BAC end page:http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 136 row: M column: 7

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..582

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-136M7"

/sex="Female"

/lab_host="DH10B"

/clone_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:

EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcoRI and EcoRI Methylase. Size

selected DNA was cloned into the pBACE3.6 vector at the

EcoRI sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies). "

ORIGIN

Query Match

Best Local Similarity

Matches 16; Conservative

90.6%; Score 15.4; DB 11; Length 582;

94.1%; Pred. No. 3.6e+03;

Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
|||
Db 185 GCACACGCTTGCCGAG 169
|||

RESULT 56

CE755799

LOCUS

DEFINITION

tigr-gss-dog-17000369661783 Dog Library Canis familiaris genomic,

genomic survey sequence.

ACCESSION CE755799

VERSION CE755799.1

KEYWORDS GSS.

SOURCE Canis familiaris (dog)

ORGANISM

Canis familiaris

REFERENCE

AUTHORS

1 (bases 1 to 586)

Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,

Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and

Venter, J.C.

TITLE The dog genome: survey sequencing and comparative analysis

JOURNAL Science 301 (5641), 1898-1903 (2003)

PUBMED 14512627

COMMENT

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

1..586

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from

peripheral blood"

ORIGIN

Query Match

Best Local Similarity

Matches 16; Conservative

90.6%; Score 15.4; DB 12; Length 586;

94.1%; Pred. No. 3.6e+03;

Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGCCGAG 17
|||
Db 10 GCCACGCTTGCCGAG 26
|||

RESULT 57

DA940956

LOCUS

DEFINITION

DA940956 SPLEN2 Homo sapiens cdna clone SPLEN2015028 5', mRNA

sequence.

ACCESSION DA940956

VERSION DA940956.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 593)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,

Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,

Ihigai, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,

Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,

Seq primer: M13-21 (TGTAACACGACGGCCAGT)

POLYA=No.

FEATURES

source

Location/Qualifiers

1..601

/organism="Pinus taeda"

/mol_type="mRNA"

/strain="3 CCLONES"

/db_xref="taxon:3352"

/clone="GE01_2_F06_A029"

/lab_host="DH10B-T1 phage-resistant E. coli"

/clone_lib="Root gravitropism April 2003 test"

/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 27 days (April 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24 h) prior to harvesting roots for mRNA preparation, the potted trees were tipped 90 degree to lay the pots on their sides to induce a gravitropic response. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 90.6%; Score 15.4; DB 8; Length 601;
 Best Local Similarity 94.1%; Pred. No. 3.6e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 427 GCCCAGCGTTGGCCGAG 443

RESULT 60

Cw480200

LOCUS

DEFINITION fbb001f239c09f0 Sorghum methylation filtered library (LibID: 104)
 Sorghum bicolor genomic clone fbb001f239c09, genomic survey
 sequence.

ACCESSION Cw480200

VERSION Cw480200.1

KEYWORDS GI:55228161

SOURCE Sorghum bicolor (sorghum)

ORGANISM Sorghum bicolor

REFERENCE

AUTHORS

Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holeman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.

TITLE

JOURNAL

PUBMED

COMMENT

Sorghum genome sequencing by methylation filtration
 PLoS Biol. 3 (1), e13 (2005)

15660154

Contact: Bedell JA

Orion Genomics, LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: fbb001f239 row: c column: 09

Seq primer: f Forward

Class: methylation filtered

High quality sequence stop: 609.

Location/Qualifiers

1..609

/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/cultivar="ATx623"

/db_xref="taxon:4558"

/clone="fbb001f239c09"

/clone_lib="Sorghum methylation filtered library (LibID: 104)"

/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 90.6%; Score 15.4; DB 13; Length 609;
 Best Local Similarity 94.1%; Pred. No. 3.6e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 437 GCCCAGCGTTGGCCGAG 453

RESULT 61

BG929309

LOCUS

DEFINITION HNC60-1-H5.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
 sequence.

ACCESSION BG929309

VERSION BG929309.1

KEYWORDS GI:14323832

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 669)

Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J., Sathe, G., Mui, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and Lark, M.W.

IDENTIFICATION AND INITIAL CHARACTERIZATION OF 5000 EXPRESSED SEQUENCED TAGS (ESTs) EACH FROM ADULT HUMAN NORMAL AND OSTEOARTHRIC CARTILAGE CDNA LIBRARIES

Osteoarthr. Cartil. 9 (7), 641-653 (2001)

JOURNAL

PUBMED

COMMENT

Contact: Sanjay Kumar

UW2109

GlaxoSmithKline

709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA

Tel: 610-270-7245

Fax: 610-270-5598

Email: sanjay_kumar-lsgsk.com

Seq primer: T7

Location/Qualifiers

1..669

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="cartilage"

/lab_host="E.coli DH10 B"

/clone_lib="HNC (Human Normal Cartilage)"

/note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI; Directional"

ORIGIN

Query Match 90.6%; Score 15.4; DB 2; Length 669;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 559 GCCCAGCGTTGGCCGAG 575

```

RESULT 62
CW480201/c
LOCUS
DEFINITION
    671 bp DNA linear GSS 02-NOV-2004
    fbb001f239c09k0 Sorghum methylation filtered library (LibID: 104)
    Sorghum bicolor genomic clone fbb001f239c09, genomic survey
    sequence.
ACCESSION
CW480201
VERSION
CW480201.1 GI:55228162
KEYWORDS
GSS.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 671)
AUTHORS
Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
McMenamy, J., Smith, M., Holeman, H., Roe, B.A., Wiley, G., Korf, I.F.,
Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddloh, J.A. and
Martienssen, R.A.
TITLE
Sorghum genome sequencing by methylation filtration
JOURNAL
PLoS Biol. 3 (1), e13 (2005)
PUBMED
15660154
COMMENT
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fbb001f239 row: c column: 09
Seq primer: k Reverse
Class: methylation filtered
High quality sequence stop: 671.
FEATURES
    source
        1..671
            /organism="Sorghum bicolor"
            /mol_type="genomic DNA"
            /cultivar="ATx623"
            /db_xref="taxon:4558"
            /clone_lib="fbb001f239c09"
            /clone="fbb001f239c09"
            /note="Sorghum methylation filtered library (LibID:
            104)"
            /note="Organ: leaf; Vector: pBCK(-); Site: 1: HincII; DNA
            prepared from purified nuclei was randomly sheared,
            end-repaired, size fractionated to enrich for the 0.5 to 5
            kb fraction, ligated into HincII-digested pBCK(-) vector
            and electroporated into E. coli cells. This is a
            methylation filtered library."
ORIGIN
    Query Match 90.6%; Score 15.4; DB 13; Length 671;
    Best Local Similarity 94.1%; Pred. No. 3.7e+03;
    Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 661 GCCCAGCGTTGGCCGCG 645

RESULT 63
DN265926
LOCUS
DEFINITION
    689 bp mRNA linear EST 02-MAR-2005
    LIB30320_043_C10 SP6 1 LIB30320 Canis familiaris cDNA clone
    LIB30320_043_C10, mRNA sequence.
ACCESSION
DN265926
VERSION
DN265926.1 GI:60434536
KEYWORDS
EST.
SOURCE
Canis familiaris (dog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

```

```

REFERENCE
    1 (bases 1 to 689)
    Statens, N.R.
    Direct Submission (Statens, N.R.)
    JOURNAL
    Unpublished (2005)
    COMMENT
    Contact: Nick Statens
    Tel: 636 247 6855
    Email: nicholas.r.statens@pfizer.com.
FEATURES
    Location/Qualifiers
        source
            1..689
                /organism="Canis familiaris"
                /mol_type="mRNA"
                /db_xref="taxon:9615"
                /clone="LIB30320_043_C10"
                /tissue_type="lymph node"
                /lab_host="DH10B"
                /clone_lib="LIB30320"
                /note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
                barcode 459"
ORIGIN
    Query Match 90.6%; Score 15.4; DB 9; Length 689;
    Best Local Similarity 94.1%; Pred. No. 3.7e+03;
    Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 612 GCCCAGCGTTGGCCGAG 628

RESULT 64
DN441215
LOCUS
DEFINITION
    694 bp mRNA linear EST 08-MAR-2005
    LIB5338-114-Al-K2-B9 LIB5338 Canis familiaris cDNA clone
    CLN14242884, mRNA sequence.
ACCESSION
DN441215
VERSION
DN441215.1 GI:60637460
KEYWORDS
EST.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
    1 (bases 1 to 694)
    Statens, N.R.
    Direct Submission (Statens, N.R.)
    JOURNAL
    Unpublished (2005)
    COMMENT
    Contact: Nick Statens
    Tel: 636 247 6855
    Email: nicholas.r.statens@pfizer.com.
FEATURES
    Location/Qualifiers
        source
            1..694
                /organism="Canis familiaris"
                /mol_type="mRNA"
                /db_xref="taxon:9615"
                /clone="CLN14242884"
                /tissue_type="kidney"
                /lab_host="DH10B"
                /clone_lib="LIB5338"
                /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; whole
                organ"
ORIGIN
    Query Match 90.6%; Score 15.4; DB 9; Length 694;
    Best Local Similarity 94.1%; Pred. No. 3.7e+03;
    Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 276 GCCCAGCGTTGGCCGAG 292

RESULT 65
COL96932

```


LOCUS
 DEFINITION GE01_2_E06_g1_A029 Root gravitropism April 2003 test Pinus taeda linear mRNA EST 21-JUN-2004
 ACCESSION cDNA clone GE01_2_E06_A029 5', mRNA sequence.
 VERSION CO196932
 KEYWORDS CO196932.1 GI:49008107
 SOURCE EST.
 ORGANISM Pinus taeda (loblolly pine)
 Pinus taeda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 REFERENCE 1 (bases 1 to 721)
 AUTHORS Pratt,J.F., Cordonnier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and Dean,J.F.D.
 TITLE An EST database from gravitropically stimulated loblolly pine
 JOURNAL (Pinus taeda) roots
 COMMENT Unpublished (2004)
 Other ESTs: GE01_2_E06_b1_A029
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
 Seq primer: JENREV (CAGAACACGTATGACC).
 FEATURES
 source Location/Qualifiers
 1..721
 /organism="Pinus taeda"
 /mol_type="mRNA"
 /strain="3 CCLONES"
 /db_xref="taxon:3352"
 /clone="GE01_2_E06_A029"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Root gravitropism April 2003 test"
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 27 days (April 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24 h) prior to harvesting roots for mRNA preparation, the potted trees were tipped 90 degree to lay the pots on their sides to induce a gravitropic response. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."
 ORIGIN
 Query Match 90.6%; Score 15.4; DB 8; Length 721;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GCCACGCTTGCCGAG 17
 |||||
 Db 567 GCCACGCTTGCCGAG 583
 RESULT 66
 CC111772/c
 LOCUS CC111772 744 bp DNA linear GSS 16-APR-2003

DEFINITION NDL.46M15.T7 Notre Dame Liverpool Aedes aegypti genomic clone
 ACCESSION NDL.46M15, genomic survey sequence.
 VERSION CC111772
 KEYWORDS CC111772.1 GI:29980827
 SOURCE GSS.
 ORGANISM Aedes aegypti (yellow fever mosquito)
 Aedes aegypti
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
 REFERENCE 1 (bases 1 to 744)
 AUTHORS Loftus,B., Shetty,J., Knudson,D. and Severson,D.
 TITLE BAC end sequencing of Aedes aegypti
 JOURNAL Unpublished (2003)
 COMMENT Other GSSs: NDL.46M15.SP6
 Contact: Brendan Loftus
 Department of Eukaryotic Genomics
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: enta@tigr.org
 Library was provided by David Severson
 Seq primer: T7
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1..744
 /organism="Aedes aegypti"
 /mol_type="genomic DNA"
 /strain="Liverpool"
 /db_xref="taxon:7159"
 /clone="NDL.46M15"
 /clone_lib="Notre Dame Liverpool"
 /note="Vector: pSCBAC1; Site1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"
 ORIGIN
 Query Match 90.6%; Score 15.4; DB 12; Length 744;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GCCACGCTTGCCGAG 17
 |||||
 Db 138 GCCACGCTTGCCGAG 122
 RESULT 67
 CW576188
 LOCUS CW576188 749 bp DNA linear GSS 22-OCT-2004
 DEFINITION OA_ABA0104L02.f OA_ABA Oryza australiensis genomic clone
 ACCESSION OA_ABA0104L02 5', genomic survey sequence.
 VERSION CW576188
 KEYWORDS CW576188.1 GI:54458858
 SOURCE GSS.
 ORGANISM Oryza australiensis
 Oryza australiensis
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 749)
 AUTHORS Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
 TITLE OMAP Project
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG
 BACKWARD: CAC TCA TTA GGC ACC CCA
 Plate: 0104 row: L column: 02
 Seq primer: TAA TAC GAC TCA CTA TAG GG
 Class: BAC ends.

FEATURES

Location/Qualifiers

1..749
 /organism="Oryza australiensis"
 /mol_type="genomic DNA"
 /db_xref="taxon:4532"
 /clone="OA_Aba0104L02"
 /tissue_type="young leaves"
 /lab_host="DH10B T1 phage resistant"
 /clone_lib="OA_ABA"
 /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 90.6%; Score 15.4; DB 13; Length 749;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17

Db 379 GCTACACGCTGGCCGAG 395

RESULT 68

CR374717

LOCUS CR374717 788 bp mRNA linear EST 22-APR-2004
 DEFINITION CR374717 AGENAE Rainbow trout normalized testis library (tcbl)
 Oncorhynchus mykiss cDNA clone tcb10013c.f.10 5prim, mRNA sequence.

ACCESSION CR374717

VERSION CR374717.1 GI:46500282

KEYWORDS EST

SOURCE Oncorhynchus mykiss (rainbow trout)

ORGANISM

Oncorhynchus mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (bases 1 to 788)

Goveroun, M., Guiguen, Y. and Le Gac, F.

Construction and primary characterization of normalized cDNA

libraries in rainbow trout, *Oncorhynchus mykiss*

Unpublished (2003)

Contact: Guiguen Y

INRA - SCRIBE

Campus de beaulieu, RENNES cedex, 35042, France

Tel: 02.23.48.50.09

Fax: 02.23.48.50.20

Email: Yann.Guiguen@beaulieu.rennes.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us
 at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
 sequence.

Plate: 0013 row: f column: 10

Seq primer: M13R.

Location/Qualifiers

1..788

/organism="Oncorhynchus mykiss"

/mol_type="mRNA"

/db_xref="taxon:8022"

/clone="tcb10013c.f.10"

/tissue_type="testis"

/lab_host="DH10B"

/clone_lib="AGENAE Rainbow trout normalized testis library

(tcbl)"

/note="Vector: p7T3D-PacI; Clone distribution : AGENAE

Resource centre. Francois PIUMI,

Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et

Etude du genome (LUREG), Domaine de Vilvert, 78352,

Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33

(0) 1.34.65.22.73"

ORIGIN

Query Match 90.6%; Score 15.4; DB 8; Length 788;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17

Db 341 GCCACGCTGGCCGAG 357

RESULT 69

BF700990/c

LOCUS BF700990/c

DEFINITION 602127938F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4284825 5',
 mRNA sequence.

ACCESSION BF700990

VERSION BF700990.1 GI:11986315

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

1 (bases 1 to 797)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-t@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCm118 row: 1 column: 10

High quality sequence start: 4

High quality sequence stop: 500.

FEATURES

source

1..797

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4284825"

/tissue_type="primitive neuroectoderm"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC 56"

/note="Organ: brain; Vector: pDNR-LTB (Clontech); Site_1:

SfiI (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc);

Double-stranded cDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCGCGCGGCACATG-dT(30)BN-3',

(where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 90.6%; Score 15.4; DB 7; Length 797;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17

Db 673 GCCACGCTGGCCGAG 657

RESULT 70

CO196873

LOCUS CO196873

798 bp mRNA linear EST 21-JUN-2004

DEFINITION GE01_2_E06.bl A029 Root gravitropism April 2003 test Pinus taeda cDNA clone GE01_2_E06_A029 3', mRNA sequence.
ACCESSION CO196873
VERSION CO196873.1 GI:49008048
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 798)
AUTHORS Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and Dean,J.F.D.
TITLE An EST database from gravitropically stimulated loblolly pine (Pinus taeda) roots
JOURNAL Unpublished (2004)
COMMENT Other ESTs: GE01_2_E06.g1_A029
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
 Seq primer: M13-21 (TGTAACGACGCGCCAGT)
 POLYA=No.

FEATURES source
 Location/Qualifiers
 1..798
 /organism="Pinus taeda"
 /mol_type="mRNA"
 /strain="3 CCLONES"
 /db_xref="taxon:3352"
 /clone="GE01_2_E06_A029"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Root gravitropism April 2003 test"
 /notes="Vector: pSL180; Site:1: EcoRI; Site:2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 27 days (April 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24 h) prior to harvesting roots for mRNA preparation, the potted trees were tipped 90 degree to lay the pots on their sides to induce a gravitropic response. Double-stranded cDNA was cloned unidirectionally into pSL180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."
ORIGIN
 Query Match 90.6%; Score 15.4; DB 8; Length 798;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCCCAGCTTGCCGAG 17
 Db 578 GCCCAGCTTGCCGAG 594
 RESULT 71
 BZ455483/c
 LOCUS 802 bp DNA linear GSS 13-DEC-2002

DEFINITION BONIS66TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONIS66, genomic survey sequence.
ACCESSION BZ455483
VERSION BZ455483.1 GI:26732690
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 802)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Uterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other GSSs: BONIS66TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
FEATURES source
 Location/Qualifiers
 1..802
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TOL000H3"
 /db_xref="taxon:3712"
 /clone="BONIS66"
 /clone_lib="BO.1.6.2 KB tot"
 /note="Vector: pHOS1; Site 1: BstXI; 1.6-2 Kb sheared total DNA inserted into pHOS1 using BstXI linkers"
ORIGIN
 Query Match 90.6%; Score 15.4; DB 11; Length 802;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GCCCAGCTTGCCGAG 17
 Db 741 GCCCAGCTTGCCGAG 725
 RESULT 72
 CW616903
 LOCUS 816 bp DNA linear GSS 22-OCT-2004
DEFINITION OA.Aba0161E12.f OA.Aba Oryza australiensis genomic clone
 OA.Aba0161E12 5', genomic survey sequence.
ACCESSION CW616903
VERSION CW616903.1 GI:54538590
KEYWORDS GSS.
SOURCE Oryza australiensis
ORGANISM Oryza australiensis
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzaceae; Oryza. 1 (bases 1 to 816)
AUTHORS Kim,H., Yu,X., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu

```

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0161 row: E column: 12
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..816
     organism="Oryza australiensis"
     /mol_type="genomic DNA"
     /db_xref="taxon:4532"
     /clone="OA_Aba0161E12"
     /tissue_type="young leaves"
     /lab_host="DH10B T1 phage resistant"
     /clone_lib="OA_Aba"
     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match          90.6%; Score 15.4; DB 13; Length 816;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAG 17
Db 320 GCTCAGCGTTGGCGGAG 336

RESULT 73
LOCUS CW646415
DEFINITION OA_Aba0185C05.f OA_Aba Oryza australiensis genomic clone
ACCESSION CW646415
VERSION CW646415.1 GI:54686385
KEYWORDS GSS.
SOURCE Oryza australiensis
ORGANISM Oryza australiensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 830)
Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0185 row: C column: 05
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..830
     organism="Oryza australiensis"
     /mol_type="genomic DNA"
     /db_xref="taxon:4532"
     /clone="OA_Aba0185C05"
     /tissue_type="young leaves"
     /lab_host="DH10B T1 phage resistant"
     /clone_lib="OA_Aba"
     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match          90.6%; Score 15.4; DB 13; Length 830;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGAG 17
Db 347 GCCACGCTTGGCGGAG 331

RESULT 75
LOCUS CC112512/c
DEFINITION NDL.37C12.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
ACCESSION CC112512
VERSION CC112512.1 GI:29981567
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)

```

```

Qy 1 GCCACGCTTGGCGGAG 17
Db 379 GCTCAGCGTTGGCGGAG 395

RESULT 74
LOCUS CA495187/c
DEFINITION AGENCOURT 10812602 NIH MGC 154 Mus musculus cDNA clone
IMAGE:6776333 5', mRNA_sequence.
ACCESSION CA495187
VERSION CA495187.1 GI:24958247
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 851)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Bradfield Laboratory
cDNA Library Preparation: Mark Bittinger
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LKMI50 row: m column: 04
High quality sequence stop: 230.

FEATURES             Location/Qualifiers
     source            1..851
     organism="Mus musculus"
     /mol_type="mRNA"
     /db_xref="taxon:10090"
     /clone="IMAGE:6776333"
     /lab_host="DH10B (phage-resistant)"
     /note="Organ: Kidney; TCDD (DMSO vehicle) Treated 48
Hours IP injections; Vector: pDONR201; Site_1: attP2;
Site_2: attP1; cDNA made by oligo-dr with attB2 site and
directionally cloned. Priming sequence:
5'-TTTCTGTCAGCGCGCCACCTTTGTACAGAAAGCTGGGTTTTTTTTTTT
TTTT-3'. Full-length enriched library was constructed
using the GeneRacer kit by Invitrogen, library
amplification 16 cycles. Library constructed by Mark
Bittinger in the Bradfield laboratory (McArdle Laboratory
for Cancer Research, University of Wisconsin). Note: this
is a NIH_MGC Library."

ORIGIN
Query Match          90.6%; Score 15.4; DB 4; Length 851;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCGGAG 17
Db 347 GCCACGCTTGGCGGAG 331

RESULT 75
LOCUS CC112512/c
DEFINITION NDL.37C12.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
ACCESSION CC112512
VERSION CC112512.1 GI:29981567
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)

```

```

ORGANISM      Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea,
Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
REFERENCE
AUTHORS      Loftus,B., Shetty,J., Knudson,D. and Severson,D.
TITLE        BAC end sequencing of Aedes aegypti
JOURNAL      Unpublished (2003)
COMMENT      Other GSSs: NDL.37C12.T7
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: SP6
Class: BAC ends.
FEATURES
source
location/Qualifiers
1..899
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/db_xref="taxon:7159"
/clone="NDL.37C12"
/note="Vector: pECBAC1; Site_1: Hind III; The library was
prepared from whole body tissue of newly hatched L1 larvae
by David Severson at the University of Notre Dame and
Hongbin Zhang"
ORIGIN
Query Match      90.6%; Score 15.4; DB 12; Length 899;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCTTGCCGAG 17
|||||
Db 152 GCCCAGCATGCCGAG 136

RESULT 76
CC851671/c
LOCUS        CC851671 914 bp DNA linear GSS 24-JUL-2003
DEFINITION  NDL.55M12.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
Notrebame Liverpool-55M12, genomic survey sequence.
ACCESSION   CC851671
VERSION     CC851671.1 GI:33203800
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
ORGANISM    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
REFERENCE
AUTHORS      Loftus,B., Shetty,J., Knudson,D. and Severson,D.
TITLE        BAC end sequencing of Aedes aegypti
JOURNAL      Unpublished (2003)
COMMENT      Other GSSs: NDL.55M12.T7
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: SP6
Class: BAC ends.
FEATURES
source
location/Qualifiers
1..914
/organism="Aedes aegypti"

```

```

/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="Notrebame Liverpool-55M12"
/clone_lib="Notre Dame Liverpool"
/note="Vector: pECBAC1; Site_1: Hind III; The library was
prepared from whole body tissue of newly hatched L1 larvae
by David Severson at the University of Notre Dame and
Hongbin Zhang"
ORIGIN
Query Match      90.6%; Score 15.4; DB 12; Length 914;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCTTGCCGAG 17
|||||
Db 189 GCCCAGCATGCCGAG 173

RESULT 77
CNS02J9X
LOCUS        CNS02J9X 929 bp DNA linear GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
143K05 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION   AL199950
VERSION     AL199950.1 GI:7838101
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE
AUTHORS      Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quettier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
10835645
2
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quettier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
10899143
3 (bases 1 to 929)
Direct Submission
Genoscope.
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
location/Qualifiers
1..929
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="143K05"
/clone_lib="G"
/note="Genoscope sequence ID : COAG143AF03SP1
end : PUC-Ori"
ORIGIN
Query Match      90.6%; Score 15.4; DB 14; Length 929;

```

```

Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCGGAG 17
Db 448 GCCCAGCGTTGGCGGAG 464

RESULT 78
LOCUS BQ685729
DEFINITION AGENCOURT_8344399 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251057
5', mRNA sequence.
ACCESSION BQ685729
VERSION BQ685729.1 GI:21811045
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LICM2395 row: b column: 18
High quality sequence stop: 538.
FEATURES
source
Location/Qualifiers
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6251057"
/tissue_type="ductal carcinoma, cell line"
/lab_hosts="NIH MGC 110"
/clone_lib="NIH MGC 110"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 90.6%; Score 15.4; DB 3; Length 940;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCGGAG 17
Db 687 GCCCAGCGTTGGCGGAG 703

RESULT 79
LOCUS DW341470
DEFINITION PP LEC0004H16f Peach shoot Prunus persica cDNA clone
PP LEC0004H16f, mRNA sequence.
ACCESSION DW341470
VERSION DW341470.1 GI:84627860
KEYWORDS EST.
SOURCE Prunus persica (peach)

```

```

Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
1 (bases 1 to 1058)
Abbott, A., Zhebentyayeva, T., Main, D., Jung, S., Staton, M.,
Jesudurai, C. and Wing, R.
Peach Model Genome for Rosaceae Shoot ESTs
Unpublished (2006)
Contact: Abbott, A.
Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aabbott@clemson.edu
Total High Quality bases = 109
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 1058.
FEATURES
source
Location/Qualifiers
1..1058
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEC0004H16f"
/tissue_type="Shoot"
/lab_host="E. coli"
/clone_lib="Peach shoot"
/notes="Vector: pBluescript II SK(-); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 90.6%; Score 15.4; DB 10; Length 1058;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCGGAG 17
Db 968 GCCCAGCGTTGGCGGAG 984

RESULT 80
LOCUS DR125588/c
DEFINITION 49128899 Drosophila pseudoobscura embryonic cDNA library Drosophila
pseudoobscura cDNA clone G11 5', mRNA sequence.
ACCESSION DR125588
VERSION DR125588.1 GI:67842286
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1117)
Richards, S., Liu, Y., Bettencourt, B.R., Hradycky, P., Letovsky, S.,
Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Meisel, R.P.,
Couronne, O., Hua, S., Smith, M.A., Zhang, P., Liu, J., Bussemaker, H.J.,
van Batenburg, M.F., Howells, S.L., Scherer, S.E., Sodergren, E.,
Matthews, B.B., Crosby, M.A., Schroeder, A.J., Ortiz-Barrientos, D.,
Rives, C.M., Metzker, M.L., Muzny, D.M., Scott, G., Steffen, D.,
Wheeler, D.A., Worley, K.C., Havlak, P., Durbin, K.J., Sgan, A.,
Gill, R., Hume, J., Morgan, M.B., Miner, G., Hamilton, C., Huang, Y.,
Waldron, L., Verduzco, D., Clerc-Blankenburg, K.P., Dubchak, I.,
Noor, M.A., Anderson, W., White, K.P., Clark, A.G., Schaeffer, S.W.,
Gelbart, W., Weinstock, G.M. and Gibbs, R.A.
Comparative genome sequencing of Drosophila pseudoobscura:
chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085
Contact: Stephen Richards
Human Genome Sequencing Center

```

Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephenr@bcm.tmc.edu
NCBI Trace Archive: 226701216
Insert Length: 1750 Std Error: 0.25.

FEATURES

source

Location/Qualifiers
1. .1117
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="G11"
/dev_stage="0-18h embryos"
/clone_lib="Drosophila pseudoobscura embryonic cDNA library"
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo dt priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match 90.6%; Score 15.4; DB 9; Length 1117;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

Db 1009 GCTCAGCTTGGCCGAG 993

RESULT 81

BE729821/c

LOCUS

DEFINITION 601564735F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3926001 5',
mRNA sequence.

ACCESSION BE729821

VERSION BE729821.1 GI:10143813

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

1 (bases 1 to 1156)

NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM753 row: e column: 10
High quality sequence start: 16
High quality sequence stop: 545.

FEATURES

source

Location/Qualifiers
1. .1156
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3926001"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_20"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 90.6%; Score 15.4; DB 7; Length 1156;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

Db 96 GCCCATGCTTGGCCGAG 80

RESULT 82

CD256036

LOCUS

DEFINITION AGENCOURT 14159547 NICHDXGC Brnl Xenopus laevis cDNA clone
IMAGE:6952009 5', mRNA sequence.

ACCESSION CD256036

VERSION CD256036.1

KEYWORDS GI:31016502

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE

1 (bases 1 to 1263)

NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov

Tissue Procurement:

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM14571 row: p column: 24

High quality sequence stop: 425.

FEATURES

source

Location/Qualifiers
1. .1263
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6952009"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC_Brnl"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."

ORIGIN

Query Match 90.6%; Score 15.4; DB 5; Length 1263;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17

Db 620 GCCCATCTTGGCCGAG 636

RESULT 83

CNS05ZL3

LOCUS

DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CNS05ZL3

VERSION CR648801.2

GI:56244151


```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1
Jailion,O., Aury,J.M., Brunet,F., Petit,J.L., Stange-Thomann,N.,
Mauceli,E., Bouneau,L., Fischer,C., Ozouf-Costaz,C., Bernot,A.,
Nicaud,S., Jaffe,D., Fisher,S., Lutfalla,G., Dossat,C.,
Segurens,B., Dasilva,C., Salanoubat,M., Levy,M., Boudet,N.,
Castellano,S., Anthouard,V., Jubin,C., Castellini,V., Katinka,M.,
Vacherie,B., Biemont,C., Skalli,Z., Cattolico,L., Poulain,J., De
Berardinis,V., Cruaud,C., Duprat,S., Brottier,P., Coutanceau,J.P.,
Gouzy,J., Parra,G., Lardier,G., Chapple,C., McKernan,K.J.,
McEwan,P., Bosak,S., Kellis,M., Volff,J.N., Guigo,R., Zody,M.C.,
Mesirov,J., Lindblad-Toh,K., Lautet,V., Schachter,V., Quetier,F.,
Robinson-Rechavi,M., Laudet,V., Schachter,V., Quetier,F.,
Saurin,W., Scarpelli,C., Wincker,P., Lander,E.S., Weissbach,J.,
and Roest Crolius,H.
Genome duplication in the teleost fish Tetraodon nigroviridis
reveals the early vertebrate proto-karyotype
Nature 431 (7011), 946-957 (2004)
15496914
PUBMED
REFERENCE
2 (bases 1 to 1388)
Genoscope.
Direct Submission
Submitted (25-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51145246.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source
1. .1388
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Query Match 90.6%; Score 15.4; DB 6; Length 1388;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17
|||||
Db 707 GCCACGCTAGCCGAG 723

RESULT 84
CNSOGOND 1454 bp mRNA linear HTC 05-JUL-2005
LOCUS
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR696861
VERSION CR696861.2 GI:56287441
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1
Jailion,O., Aury,J.M., Brunet,F., Petit,J.L., Stange-Thomann,N.,
Mauceli,E., Bouneau,L., Fischer,C., Ozouf-Costaz,C., Bernot,A.,
Nicaud,S., Jaffe,D., Fisher,S., Lutfalla,G., Dossat,C.,
Segurens,B., Dasilva,C., Salanoubat,M., Levy,M., Boudet,N.,
Castellano,S., Anthouard,V., Jubin,C., Castellini,V., Katinka,M.,
Vacherie,B., Biemont,C., Skalli,Z., Cattolico,L., Poulain,J., De
Berardinis,V., Cruaud,C., Duprat,S., Brottier,P., Coutanceau,J.P.,
Gouzy,J., Parra,G., Lardier,G., Chapple,C., McKernan,K.J.,
McEwan,P., Bosak,S., Kellis,M., Volff,J.N., Guigo,R., Zody,M.C.,
Mesirov,J., Lindblad-Toh,K., Lautet,V., Schachter,V., Quetier,F.,
Robinson-Rechavi,M., Laudet,V., Schachter,V., Quetier,F.,
Saurin,W., Scarpelli,C., Wincker,P., Lander,E.S., Weissbach,J.,
and Roest Crolius,H.
Genome duplication in the teleost fish Tetraodon nigroviridis
reveals the early vertebrate proto-karyotype
Nature 431 (7011), 946-957 (2004)
15496914
PUBMED
REFERENCE
2 (bases 1 to 1388)
Genoscope.
Direct Submission
Submitted (25-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51145246.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source
1. .1388
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Query Match 90.6%; Score 15.4; DB 6; Length 1388;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17
|||||
Db 707 GCCACGCTAGCCGAG 723

RESULT 84
CNSOGOND 1454 bp mRNA linear HTC 05-JUL-2005
LOCUS
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR696861
VERSION CR696861.2 GI:56287441
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1
Jailion,O., Aury,J.M., Brunet,F., Petit,J.L., Stange-Thomann,N.,
Mauceli,E., Bouneau,L., Fischer,C., Ozouf-Costaz,C., Bernot,A.,
Nicaud,S., Jaffe,D., Fisher,S., Lutfalla,G., Dossat,C.,
Segurens,B., Dasilva,C., Salanoubat,M., Levy,M., Boudet,N.,
Castellano,S., Anthouard,V., Jubin,C., Castellini,V., Katinka,M.,
Vacherie,B., Biemont,C., Skalli,Z., Cattolico,L., Poulain,J., De
Berardinis,V., Cruaud,C., Duprat,S., Brottier,P., Coutanceau,J.P.,
Gouzy,J., Parra,G., Lardier,G., Chapple,C., McKernan,K.J.,
McEwan,P., Bosak,S., Kellis,M., Volff,J.N., Guigo,R., Zody,M.C.,
Mesirov,J., Lindblad-Toh,K., Lautet,V., Schachter,V., Quetier,F.,
Robinson-Rechavi,M., Laudet,V., Schachter,V., Quetier,F.,
Saurin,W., Scarpelli,C., Wincker,P., Lander,E.S., Weissbach,J.,
and Roest Crolius,H.
Genome duplication in the teleost fish Tetraodon nigroviridis
reveals the early vertebrate proto-karyotype
Nature 431 (7011), 946-957 (2004)
15496914
PUBMED
REFERENCE
2 (bases 1 to 1388)
Genoscope.
Direct Submission
Submitted (25-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51145246.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source
1. .1388
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Query Match 90.6%; Score 15.4; DB 6; Length 1388;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCACGCTGGCCGAG 17
|||||
Db 707 GCCACGCTAGCCGAG 723

RESULT 85
CNSOFYVC 1492 bp mRNA linear HTC 05-JUL-2005
LOCUS
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR694556
VERSION CR694556.2 GI:56286437
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1
Jailion,O., Aury,J.M., Brunet,F., Petit,J.L., Stange-Thomann,N.,
Mauceli,E., Bouneau,L., Fischer,C., Ozouf-Costaz,C., Bernot,A.,
Nicaud,S., Jaffe,D., Fisher,S., Lutfalla,G., Dossat,C.,
Segurens,B., Dasilva,C., Salanoubat,M., Levy,M., Boudet,N.,
Castellano,S., Anthouard,V., Jubin,C., Castellini,V., Katinka,M.,
Vacherie,B., Biemont,C., Skalli,Z., Cattolico,L., Poulain,J., De
Berardinis,V., Cruaud,C., Duprat,S., Brottier,P., Coutanceau,J.P.,
Gouzy,J., Parra,G., Lardier,G., Chapple,C., McKernan,K.J.,
McEwan,P., Bosak,S., Kellis,M., Volff,J.N., Guigo,R., Zody,M.C.,
Mesirov,J., Lindblad-Toh,K., Lautet,V., Schachter,V., Quetier,F.,
Robinson-Rechavi,M., Laudet,V., Schachter,V., Quetier,F.,
Saurin,W., Scarpelli,C., Wincker,P., Lander,E.S., Weissbach,J.,
and Roest Crolius,H.
Genome duplication in the teleost fish Tetraodon nigroviridis
reveals the early vertebrate proto-karyotype
Nature 431 (7011), 946-957 (2004)
15496914
PUBMED
REFERENCE
2 (bases 1 to 1492)
Genoscope.
Direct Submission
Submitted (24-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

```

COMMENT (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51192463.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.

FEATURES

source
1. .1492
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Eyes"

ORIGIN

Query Match 90.6%; Score 15.4; DB 6; Length 1492;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
|||||
Db 818 GCCCAGCTAGCCGAG 834

RESULT 86

CNSOGM2S
LOCUS Tetraodon nigroviridis full-length cDNA. linear HTC 05-JUL-2005
DEFINITION
ACCESSION CR724638
VERSION CR724638.2 GI:56306113
KEYWORDS HTc; cDNA; full-length; full-length cDNA; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE

1
Jailon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N.,
Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A.,
Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C.,
Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Boudet, N.,
Castellano, S., Anthouard, V., Jubin, C., Castelli, V., Katinka, M.,
Vacherie, B., Biemont, C., Skalli, Z., Cattolico, L., Poulain, J., De
Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau, J.P.,
Gouzy, J., Parra, G., Lardier, G., Chapple, C., McKernan, K.J.,
McEwan, P., Bosak, S., Kellis, M., Volff, J.N., Guigo, R., Zody, M.C.,
Mesirov, J., Lindblad-Toh, K., Laudet, V., Schachter, V., Quetier, F.,
Robinson-Rechavi, M., Laudet, V., Schachter, V., Quetier, F.,
Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissenbach, J.
and Roest Crolius, H.
Genome duplication in the teleost fish Tetraodon nigroviridis
reveals the early vertebrate proto-karyotype
Nature 431 (7011), 946-957 (2004)
15496914
2 (bases 1 to 1505)
Genoscope.
Direct Submission
Submitted (24-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51222889.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. .1505
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

TITLE

JOURNAL
PUBMED
REFERENCE

2 (bases 1 to 1505)
Genoscope.

TITLE

Direct Submission
Submitted (24-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51222889.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. .1505
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

COMMENT

Query Match 90.6%; Score 15.4; DB 6; Length 1505;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;

FEATURES

source
1. .1505
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

ORIGIN

Query Match 90.6%; Score 15.4; DB 6; Length 1505;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCCAGCTTGGCCGAG 17
|||||
Db 822 GCCCAGCTAGCCGAG 838

RESULT 87

CNSOG3DQ
LOCUS Tetraodon nigroviridis full-length cDNA. linear HTC 05-JUL-2005
DEFINITION
ACCESSION CR700402
VERSION CR700402.2 GI:56289060
KEYWORDS HTc; cDNA; full-length; full-length cDNA; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE

1
Jailon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N.,
Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A.,
Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C.,
Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Boudet, N.,
Castellano, S., Anthouard, V., Jubin, C., Castelli, V., Katinka, M.,
Vacherie, B., Biemont, C., Skalli, Z., Cattolico, L., Poulain, J., De
Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau, J.P.,
Gouzy, J., Parra, G., Lardier, G., Chapple, C., McKernan, K.J.,
McEwan, P., Bosak, S., Kellis, M., Volff, J.N., Guigo, R., Zody, M.C.,
Mesirov, J., Lindblad-Toh, K., Laudet, V., Schachter, V., Quetier, F.,
Robinson-Rechavi, M., Laudet, V., Schachter, V., Quetier, F.,
Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissenbach, J.
and Roest Crolius, H.
Genome duplication in the teleost fish Tetraodon nigroviridis
reveals the early vertebrate proto-karyotype
Nature 431 (7011), 946-957 (2004)
15496914
2 (bases 1 to 1512)
Genoscope.
Direct Submission
Submitted (24-NOV-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
On Dec 3, 2004 this sequence version replaced gi:51198311.
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. .1512
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Eyes"

FEATURES

source

1. .1512
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Eyes"

ORIGIN

Query Match 90.6%; Score 15.4; DB 6; Length 1512;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
|||||
Db 837 GCCCAGCTAGCCGAG 853

RESULT 88

CNS0EX00
LOCUS Tetraodon nigroviridis full-length cDNA. linear HTC 05-JUL-2005
DEFINITION
ACCESSION CR646338
VERSION CR646338.2 GI:56243177
KEYWORDS HTc; cDNA; full-length; full-length cDNA; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Jaillon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N., Maudeli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C., Segurens, B., Da Silva, C., Salanoubat, M., Levy, M., Boudet, N., Castellan, S., Anthouard, V., Jubin, C., Castelli, V., Katinka, M., Vacherie, B., Blemont, C., Skalli, Z., Catolico, L., Poulain, J., De Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau, J.P., Guzy, J., Parra, G., Lardier, G., Chappelle, C., McKernan, K.J., McEwan, P., Bosak, S., Kellis, M., Volff, J.N., Guigo, R., Zody, M.C., Mesirov, J., Lindblad-Toh, K., Birren, B., Nusbaum, C., Kahn, D., Robinson-Rechavi, M., Laudet, V., Schachter, V., Quetier, F., Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissenbach, J. and Roest Crollius, H.

TITLE Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype

JOURNAL Nature 431 (7011), 946-957 (2004)

REFERENCE 2
AUTHORS (bases 1 to 1518)

TITLE Direct Submission

JOURNAL Submitted (25-NOV-2004) Genoscope - Centre National de Sequencage - 2 rue Gaston Crenieux, CP 5706 - 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT On Dec 3, 2004 this sequence version replaced gi:51142783. The sequences are based on single pass reads. More information available at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES
source Location/Qualifiers
1..1518
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Query Match 90.6%; Score 15.4; DB 6; Length 1518;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 841 GCCCAGCGTAGGCCGAG 857

RESULT 89
DQ049847
LOCUS DQ049847 1962 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens CLIP-2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ049847
VERSION DQ049847.1 GI:66903046
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1962)
Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE A scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees

JOURNAL PLoS Biol. 3 (6), E170 (2005)

REFERENCE 2
AUTHORS (bases 1 to 1962)
Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,

Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

Direct Submission
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES
source Location/Qualifiers
1..1962
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
1..1962
/genes="CLIP-2"
/locus_tag="HC17851"

ORIGIN
Query Match 90.6%; Score 15.4; DB 14; Length 1962;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
|||||
Db 1694 GCCCAGCGTTGGCCGAG 1710

RESULT 90
BC068256
LOCUS BC068256 4208 bp mRNA linear HTC 12-DEC-2005
DEFINITION Homo sapiens cDNA clone IMAGE:5295699, containing frame-shift errors.
ACCESSION BC068256
VERSION BC068256.1 GI:45946891
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4208)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viala, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fane, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

CONSRM
TITLE Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL 12477932
PUBMED 2 (bases 1 to 4208)

REFERENCE 1
AUTHORS NIH MGC Project
Direct Submission
Submitted (31-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk

Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amgobcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 91 Row: 1 Column: 14
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23510284
 This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
 1..4208

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5295699"
 /tissue_type="Testis"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"

ORIGIN

Query Match 90.6%; Score 15.4; DB 6; Length 4208;
 Best Local Similarity 94.1%; Pred. No. 3.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
 |||||
 Db 3283 GCCCAGCGTTGGCCAG 3299

RESULT 91
 BI529325 198 bp mRNA linear EST 29-AUG-2001
 LOCUS 1024097G09.yl C. reinhardtii CC-1690, normalized, Lambda Zap II
 DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI529325
 VERSION BI529325.1 GI:15369899
 KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
 Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 198)
 Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.
 ANALYSES OF THE CHLAMYDOMONAS REINHARDTII GENOME: A MODEL, UNICELLULAR SYSTEM FOR ANALYZING GENE FUNCTION AND REGULATION IN VASCULAR PLANTS. Project: 1024b

JOURNAL Unpublished (2001)
 COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Duke University
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.

FEATURES

Location/Qualifiers
 1..198

/organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
 /note="vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. POLYA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 88.2%; Score 15; DB 2; Length 198;
 Best Local Similarity 100.0%; Pred. No. 5.8e+03;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGA 16
 |||||
 Db 100 CCCACGCTTGCCGA 114

RESULT 92

AI613865/c 278 bp mRNA linear EST 21-APR-1999
 LOCUS V988d01.yl Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
 DEFINITION clone IMAGE:873025 5', mRNA sequence.

ACCESSION AI613865
 VERSION AI613865.1 GI:4623032
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 278)

REFERENCE 1
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)

JOURNAL Contact: Marra M/WashU-NCI Mouse EST Project 1999
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mousse@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:512505

This read is a RESEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the correct orientation)

Seq primer: -40RP from Gibco
 High quality sequence stop: 269.

FEATURES

Location/Qualifiers
 1..278

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:873025"
 /sex="mixed"
 /tissue_type="pooled organs"
 /dev_stage="7 day"
 /lab_host="DH10B"
 /clone_lib="Barstead mouse pooled organs MPLRB4"

/note="Organ: pooled; Vector: pTT3D-PacI; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [GTGATTTGGTACCT], digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT3 vector. Library constructed by Bob Barstead."

ORIGIN

Query Match 88.2%; Score 15; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGAG 17
|||||
Db 250 CCACGCTTGGCCGAG 236

RESULT 93

DR992377

LOCUS

DEFINITION M1r701124_91 Apple_EST_Mdlr Malus x domestica cDNA similar to gb|AF68396.1| unknown protein [Oryza sativa (japonica cultivar-group)], mRNA sequence.

ACCESSION

DR992377

VERSION

DR992377.1

KEYWORDS

EST.

SOURCE

ORGANISM

Malus x domestica

Malus x domestica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

1 (bases 1 to 309)

Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A., Aldwinckle, H., Malnoy, M., Carroll, N., Goldsbrough, P., Orvis, K., Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I., Teagareishvili, R., Kennedy, S., Waterston, R., and Wilson, R.

Apple Functional Genomics grant - NSF 0321702

Unpublished (2004)

Contact: Schuyler S. Korban

Apple Functional Genomics grant - NSF 0321702

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library material provided by S. Korban/H. Aldwinckle/ M. Malnoy

Library constructed by D.O.Gonzalez/L.Vodkin Library sequenced by Washington University Genome Sequencing Center

Original QR value of 517 was extended to value 517 (,)

This trace has been recalled with phred

Original value before phred recall for SL was 112

Original value before phred recall for SR was 519.

Location/Qualifiers

1. 309

/organism="Malus x domestica"

/mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/tissue_type="Leaves challenged with oblique banded leaf roller (Choristoneura rosaceana)"

/lab_host="DH10B ampicillin resistant"

/clone_lib="Apple EST Mdlr"

/note="Vector: pBluescript II SK (+); Site 1: XhoI; Site 2: EcoRI; Total RNA was extracted from freeze dried leaf tissue, using the method described by Wang and Vodkin (Plant Molecular Biology Reporter 12:132-145, 1994).

Poly(A) + mRNA was isolated from total RNA using the PolyATtract mRNA Isolation system III (Promega). The library was prepared using the Stratagene pBluescript II

XR cDNA library construction kit. Complementary DNA was

ORIGIN

Query Match 88.2%; Score 15; DB 10; Length 309;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGAG 17
|||||
Db 203 CCACGCTTGGCCGAG 217

RESULT 94

AV850077

LOCUS

DEFINITION AV850077 Nori Satoh unpublished cDNA library, cleavage stage embryo Ciona intestinalis cDNA clone rcic104n18 3', mRNA sequence.

ACCESSION

AV850077

VERSION

AV850077.1

KEYWORDS

EST.

SOURCE

ORGANISM

Ciona intestinalis

Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.

1 (bases 1 to 322)

Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.

Expressed genes in Ciona intestinalis

Unpublished (2000)

Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1. 322

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="rcic104n18"

/tissue_type="whole animal"

/dev_stage="cleaving embryo"

/clone_lib="Nori Satoh unpublished cDNA library, cleavage stage embryo"

ORIGIN

Query Match 88.2%; Score 15; DB 7; Length 322;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCACGCTTGGCCGA 16
|||||
Db 146 CCACGCTTGGCCGA 160

RESULT 95

CA661056

LOCUS

DEFINITION CA661056 wlm1.pk0017.f2 wlm1 Triticum aestivum cDNA clone wlm1.pk0017.f2 5' end, mRNA sequence.

ACCESSION

CA661056

VERSION

CA661056.1

KEYWORDS

EST.

```

SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Pooidae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 368)
AUTHORS    Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
TITLE      DuPont Wheat cDNA Sequence
JOURNAL    Unpublished (2002)
COMMENT    Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.tingey@usa.dupont.com
            Seq primer: M13.

FEATURES   1..368
            Location/Qualifiers
            source          1..368
                        /organism="Triticum aestivum"
                        /mol_type="mRNA"
                        /cultivar="Stephens"
                        /db_xref="taxon:4565"
                        /clone="wlm1.pk0017.f2"
                        /tissue_type="leaf"
                        /clone_lib="wlm1"
                        /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                        XhoI; Wheat (Triticum aestivum L.) seedlings 1 hr after
                        inoculation with Erysiphe graminis f. sp tritici"

ORIGIN
Query Match      88.2%; Score 15; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGAG 17
    |||||
Db 295 CCACGCTTGGCCGAG 309

RESULT 96
BW421246
LOCUS      BW421246          419 bp      mRNA      linear      EST 02-JUN-2004
DEFINITION      BW421246 Yutaka Satou unpublished cDNA library, mature adult whole
            animal Ciona intestinalis CDNA clone cima839j04 3', mRNA sequence.
ACCESSION      BW421246
VERSION        BW421246.1
KEYWORDS       EST.
SOURCE        Ciona intestinalis
ORGANISM      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1 (bases 1 to 419)
AUTHORS    Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE      Expressed genes in Ciona intestinalis (2004)
JOURNAL    Unpublished (2004)
COMMENT    Contact: Yutaka Satou
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4095
            Fax: 81-75-705-1113
            Email: yutaka@ascidian.zool.kyoto-u.ac.jp.
            Location/Qualifiers
            source          1..419
                        /organism="Ciona intestinalis"
                        /mol_type="mRNA"
                        /db_xref="taxon:7719"
                        /clone="cima839j04"
                        /tissue_type="whole animal"
                        /dev_stage="mature adult"
                        /clone_lib="Yutaka Satou unpublished cDNA library, mature

SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Pooidae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 368)
AUTHORS    Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
TITLE      DuPont Wheat cDNA Sequence
JOURNAL    Unpublished (2002)
COMMENT    Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.tingey@usa.dupont.com
            Seq primer: M13.

FEATURES   1..368
            Location/Qualifiers
            source          1..368
                        /organism="Triticum aestivum"
                        /mol_type="mRNA"
                        /cultivar="Stephens"
                        /db_xref="taxon:4565"
                        /clone="wlm1.pk0017.f2"
                        /tissue_type="leaf"
                        /clone_lib="wlm1"
                        /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                        XhoI; Wheat (Triticum aestivum L.) seedlings 1 hr after
                        inoculation with Erysiphe graminis f. sp tritici"

ORIGIN
Query Match      88.2%; Score 15; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGAG 17
    |||||
Db 295 CCACGCTTGGCCGAG 309

RESULT 96
BW421246
LOCUS      BW421246          419 bp      mRNA      linear      EST 02-JUN-2004
DEFINITION      BW421246 Yutaka Satou unpublished cDNA library, mature adult whole
            animal Ciona intestinalis CDNA clone cima839j04 3', mRNA sequence.
ACCESSION      BW421246
VERSION        BW421246.1
KEYWORDS       EST.
SOURCE        Ciona intestinalis
ORGANISM      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1 (bases 1 to 419)
AUTHORS    Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE      Expressed genes in Ciona intestinalis (2004)
JOURNAL    Unpublished (2004)
COMMENT    Contact: Yutaka Satou
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4095
            Fax: 81-75-705-1113
            Email: yutaka@ascidian.zool.kyoto-u.ac.jp.
            Location/Qualifiers
            source          1..419
                        /organism="Ciona intestinalis"
                        /mol_type="mRNA"
                        /db_xref="taxon:7719"
                        /clone="cima839j04"
                        /tissue_type="whole animal"
                        /dev_stage="mature adult"
                        /clone_lib="Yutaka Satou unpublished cDNA library, mature

```

```

ORIGIN      adult whole animal"

Query Match      88.2%; Score 15; DB 3; Length 419;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGA 16
    |||||
Db 188 CCCACGCTTGGCCGA 202

RESULT 97
CN899452
LOCUS      CN899452          438 bp      mRNA      linear      EST 04-JUN-2004
DEFINITION      010706ABAA999039HT (ABAA) Pacific Rose spur buds from ON trees
            Malus x domestica cDNA clone ABAA999039, mRNA sequence.
ACCESSION      CN899452
VERSION        CN899452.1
KEYWORDS       EST.
SOURCE        Malus x domestica
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
REFERENCE   1 (bases 1 to 438)
AUTHORS    Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
            McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
            HortResearch Apple EST Project
            Unpublished (2004)
            Contact: Gleave,A.
            Sequencing Facility
            The Horticulture and Food Research Institute of New Zealand Ltd
            120 Mc Albert Rd, Mc Albert, Auckland, New Zealand
            Tel: 00 64 09 815 4200
            Fax: 00 64 09 815 4201
            Email: est@hortresearch.co.nz.
            Location/Qualifiers
            source          1..438
                        /organism="Malus x domestica"
                        /mol_type="mRNA"
                        /db_xref="taxon:3750"
                        /clone="ABAA999039"
                        /tissue_type="Spur buds from 'ON' trees"
                        /clone_lib="(ABAA) Pacific Rose spur buds from ON trees"
                        /note="Vector: pBluescript SK(-); Library sequenced by
                        Genesis Research & Development"

ORIGIN

Query Match      88.2%; Score 15; DB 8; Length 438;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGAG 17
    |||||
Db 175 CCACGCTTGGCCGAG 189

RESULT 98
BW366787/c
LOCUS      BW366787/c          484 bp      mRNA      linear      EST 28-MAY-2004
DEFINITION      BW366787 Yutaka Satou unpublished cDNA library, mature adult whole
            animal Ciona intestinalis cDNA clone cima839j04 5', mRNA sequence.
ACCESSION      BW366787
VERSION        BW366787.1
KEYWORDS       EST.
SOURCE        Ciona intestinalis
ORGANISM      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1 (bases 1 to 484)
AUTHORS    Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE      Expressed genes in Ciona intestinalis (2004)

```

```
JOURNAL Unpublished (2004)
COMMENT Contact: Yutaka Satou
          Department of Zoology
          Kyoto University
          Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
          Tel: 81-75-753-4095
          Fax: 81-75-705-1113
          Email: yutaka@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source 1..484
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="cima839j04"
        /tissue_type="whole animal"
        /dev_stage="mature adult"
        /clone_lib="Yutaka Satou unpublished cDNA library, mature
        adult whole animal"

ORIGIN
Query Match 88.2%; Score 15; DB 3; Length 484;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGA 16
Db 232 CCCACGCTTGCCGA 218

RESULT 99
BW178352
LOCUS BW178352 492 bp mRNA linear EST 04-NOV-2002
DEFINITION BW178352 Nori Satoh unpublished cDNA library, heart Ciona
            intestinalis cDNA clone rc iht005e16 3', mRNA sequence.
ACCESSION BW178352
VERSION BW178352.1 GI:24568276
KEYWORDS EST.
SOURCE Ciona intestinalis
        ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
        Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 492)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
          Department of Zoology
          Kyoto University
          Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
          Tel: 81-75-753-4081
          Fax: 81-75-705-1113
          Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source 1..492
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="rc iht005e16"
        /tissue_type="heart"
        /clone_lib="Nori Satoh unpublished cDNA library, heart"

ORIGIN
Query Match 88.2%; Score 15; DB 3; Length 492;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGA 16
Db 227 CCCACGCTTGCCGA 241

RESULT 100
BW203028/c

JOURNAL Unpublished (2004)
COMMENT Contact: Yutaka Satou
          Department of Zoology
          Kyoto University
          Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
          Tel: 81-75-753-4095
          Fax: 81-75-705-1113
          Email: yutaka@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source 1..517
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="cic1083c24"
        /tissue_type="whole body"
        /dev_stage="cleaving embryo"
        /clone_lib="Nori Satoh unpublished cDNA library, cleaving
        embryo"

ORIGIN
Query Match 88.2%; Score 15; DB 3; Length 517;
Best Local Similarity 100.0%; Pred. No. 5.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGA 16
Db 275 CCCACGCTTGCCGA 261

Search completed: June 10, 2006, 19:37:09
Job time : 2150.8 secs

LOCUS BW203028 517 bp mRNA linear EST 01-JUN-2005
DEFINITION BW203028 Nori Satoh unpublished cDNA library, cleaving embryo Ciona
            intestinalis cDNA clone cic1083c24 5', mRNA sequence.
ACCESSION BW203028
VERSION BW203028.1 GI:24617457
KEYWORDS EST.
SOURCE Ciona intestinalis
        ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
        Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 517)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
          Department of Zoology
          Kyoto University
          Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
          Tel: 81-75-753-4081
          Fax: 81-75-705-1113
          Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source 1..517
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="cic1083c24"
        /tissue_type="whole body"
        /dev_stage="cleaving embryo"
        /clone_lib="Nori Satoh unpublished cDNA library, cleaving
        embryo"
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:21:38 ; Search time 38,1083 seconds
(without alignments)
834.695 Million cell updates/sec

Title: US-10-600-816-31

Perfect score: 17

Sequence: 1 gccacgccttgccgag 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq.*

2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq.*

3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq.*

4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq.*

5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq.*

6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq.*

7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq.*

8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq.*

9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq.*

10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	17	100.0	601	3 US-09-949-016-29173 Sequence 29173, A
2	17	100.0	601	3 US-09-949-016-126957 Sequence 126957, A
3	17	100.0	1212	3 US-09-188-930-249 Sequence 249, App
4	17	100.0	1212	3 US-09-312-283C-249 Sequence 249, App
5	17	100.0	1228	3 US-09-495-050A-223 Sequence 223, App
6	17	100.0	2289	3 US-09-949-016-3613 Sequence 3613, App
7	17	100.0	2456	3 US-09-949-016-572 Sequence 572, App
8	17	100.0	26086	3 US-09-949-016-15355 Sequence 15355, A
9	17	100.0	26238	3 US-09-949-016-12314 Sequence 12314, A
10	15.4	90.6	1527	3 US-09-252-991A-11321 Sequence 11321, A
11	15.4	90.6	4183	3 US-08-996-083-2 Sequence 2, Appli
12	15.4	90.6	4183	3 US-09-429-516-2 Sequence 2, Appli
13	14.4	84.7	1021	3 US-09-949-016-157184 Sequence 157184, A
14	14.4	84.7	1021	3 US-09-027-900-9 Sequence 9, Appli
15	14.4	84.7	1520	3 US-09-027-900-8 Sequence 8, Appli
16	14.4	84.7	1675	3 US-09-027-900-7 Sequence 7, Appli
17	14.4	84.7	42242	3 US-09-949-016-16127 Sequence 16127, A
18	14.4	84.7	47698	3 US-09-949-016-13514 Sequence 13514, A
19	14.4	84.7	49312	3 US-09-671-317-485 Sequence 485, App
20	14.4	84.7	67002	3 US-09-949-016-16803 Sequence 16803, A
21	13.8	81.2	20	3 US-09-332-522E-17 Sequence 17, Appl
22	13.8	81.2	47	3 US-09-422-978-2426 Sequence 2426, Ap
23	13.8	81.2	283	3 US-09-313-294A-2571 Sequence 2571, Ap

24	13.8	81.2	293	3 US-09-313-294A-5109 Sequence 5109, Ap
25	13.8	81.2	385	3 US-08-990-823-44 Sequence 44, Appl
26	13.8	81.2	385	3 US-09-477-135A-44 Sequence 44, Appl
27	13.8	81.2	600	4 US-10-233-942-3 Sequence 3, Appli
28	13.8	81.2	600	5 US-09-945-952A-3 Sequence 3, Appli
29	13.8	81.2	601	3 US-09-949-016-121284 Sequence 121284, A
30	13.8	81.2	601	3 US-09-949-016-205926 Sequence 205926, A
31	13.8	81.2	601	3 US-09-949-016-205927 Sequence 205927, A
32	13.8	81.2	795	3 US-09-902-540-5464 Sequence 5464, Ap
33	13.8	81.2	999	3 US-09-252-991A-6261 Sequence 6261, Ap
34	13.8	81.2	1176	3 US-09-252-991A-547 Sequence 547, App
35	13.8	81.2	1182	3 US-09-540-236-266 Sequence 266, App
36	13.8	81.2	1235	2 US-08-095-726-13 Sequence 13, Appl
37	13.8	81.2	1235	2 US-08-095-726-15 Sequence 15, Appl
38	13.8	81.2	1235	2 US-08-096-623A-13 Sequence 13, Appl
39	13.8	81.2	1235	2 US-08-096-623A-15 Sequence 15, Appl
40	13.8	81.2	1251	3 US-09-252-991A-1019 Sequence 1019, Ap
41	13.8	81.2	1578	2 US-08-681-129-1 Sequence 1, Appli
42	13.8	81.2	1608	3 US-09-252-991A-569 Sequence 569, App
43	13.8	81.2	1616	3 US-09-869-588-32 Sequence 32, Appl
44	13.8	81.2	1653	3 US-09-252-991A-1111 Sequence 1111, Ap
45	13.8	81.2	1976	3 US-09-165-042-2 Sequence 2, Appli
46	13.8	81.2	1983	3 US-09-252-991A-531 Sequence 531, App
47	13.8	81.2	2049	3 US-09-252-991A-6426 Sequence 6426, Ap
48	13.8	81.2	2058	2 US-08-358-117-1 Sequence 1, Appli
49	13.8	81.2	2058	3 US-08-470-588-1 Sequence 1, Appli
50	13.8	81.2	2169	3 US-09-976-594-315 Sequence 315, App
51	13.8	81.2	2174	3 US-09-613-444-1 Sequence 1, Appli
52	13.8	81.2	2174	3 US-10-142-516-1 Sequence 1, Appli
53	13.8	81.2	2184	3 US-09-902-540-5116 Sequence 5116, Ap
54	13.8	81.2	2847	3 US-09-252-991A-1036 Sequence 1036, Ap
55	13.8	81.2	3133	2 US-08-162-809-1 Sequence 1, Appli
56	13.8	81.2	3453	3 US-09-902-540-4713 Sequence 4713, Ap
57	13.8	81.2	3636	4 US-10-094-749-957 Sequence 957, App
58	13.8	81.2	3814	3 US-09-919-039-254 Sequence 254, App
59	13.8	81.2	4563	3 US-09-252-991A-930 Sequence 930, App
60	13.8	81.2	5382	3 US-09-479-122-21 Sequence 21, Appl
61	13.8	81.2	5382	3 US-09-484-997-21 Sequence 21, Appl
62	13.8	81.2	5382	3 US-09-481-355-21 Sequence 21, Appl
63	13.8	81.2	5382	3 US-09-481-282-21 Sequence 21, Appl
64	13.8	81.2	5382	3 US-09-455-659A-21 Sequence 21, Appl
65	13.8	81.2	5382	3 US-09-484-996-21 Sequence 21, Appl
66	13.8	81.2	5382	3 US-09-479-123-21 Sequence 21, Appl
67	13.8	81.2	5382	3 US-09-484-317A-21 Sequence 21, Appl
68	13.8	81.2	5382	3 US-09-276-820A-21 Sequence 21, Appl
69	13.8	81.2	5382	5 US-09-760-897A-21 Sequence 21, Appl
70	13.8	81.2	8540	3 US-08-487-283A-4 Sequence 4, Appli
71	13.8	81.2	8540	7 PCT-US96-05611A-12 Sequence 12, Appl
72	13.8	81.2	9737	3 US-09-479-122-22 Sequence 22, Appl
73	13.8	81.2	9737	3 US-09-479-122-23 Sequence 23, Appl
74	13.8	81.2	9737	3 US-09-484-997-22 Sequence 22, Appl
75	13.8	81.2	9737	3 US-09-484-997-23 Sequence 23, Appl
76	13.8	81.2	9737	3 US-09-484-997-28 Sequence 28, Appl
77	13.8	81.2	9737	3 US-09-484-997-28 Sequence 28, Appl
78	13.8	81.2	9737	3 US-09-481-355-22 Sequence 22, Appl
79	13.8	81.2	9737	3 US-09-481-355-23 Sequence 23, Appl
80	13.8	81.2	9737	3 US-09-481-355-28 Sequence 28, Appl
81	13.8	81.2	9737	3 US-09-481-282-22 Sequence 22, Appl
82	13.8	81.2	9737	3 US-09-481-282-23 Sequence 23, Appl
83	13.8	81.2	9737	3 US-09-481-282-28 Sequence 28, Appl
84	13.8	81.2	9737	3 US-09-455-659A-22 Sequence 22, Appl
85	13.8	81.2	9737	3 US-09-455-659A-23 Sequence 23, Appl
86	13.8	81.2	9737	3 US-09-455-659A-28 Sequence 28, Appl
87	13.8	81.2	9737	3 US-09-484-996-22 Sequence 22, Appl
88	13.8	81.2	9737	3 US-09-484-996-23 Sequence 23, Appl
89	13.8	81.2	9737	3 US-09-484-996-28 Sequence 28, Appl
90	13.8	81.2	9737	3 US-09-479-123-22 Sequence 22, Appl
91	13.8	81.2	9737	3 US-09-479-123-23 Sequence 23, Appl
92	13.8	81.2	9737	3 US-09-479-123-28 Sequence 28, Appl
93	13.8	81.2	9737	3 US-09-484-317A-22 Sequence 22, Appl
94	13.8	81.2	9737	3 US-09-484-317A-23 Sequence 23, Appl
95	13.8	81.2	9737	3 US-09-484-317A-28 Sequence 28, Appl
96	13.8	81.2	9737	3 US-09-276-820A-22 Sequence 22, Appl

97	13.8	81.2	9737	3	US-09-276-820A-23	Sequence 23, Appl	c 170	13.4	78.8	1548	3	US-09-545-586-11	Sequence 11, Appl
98	13.8	81.2	9737	3	US-09-276-820A-28	Sequence 28, Appl	c 171	13.4	78.8	1548	3	US-09-769-864-11	Sequence 11, Appl
99	13.8	81.2	9737	5	US-09-760-897A-22	Sequence 22, Appl	c 172	13.4	78.8	1548	3	US-10-025-648-6	Sequence 6, Appl
100	13.8	81.2	9737	5	US-09-760-897A-23	Sequence 23, Appl	c 173	13.4	78.8	1548	3	US-09-441-313-11	Sequence 11, Appl
101	13.8	81.2	9737	5	US-09-760-897A-28	Sequence 28, Appl	174	13.4	78.8	1549	3	US-09-438-938-8	Sequence 8, Appl
102	13.8	81.2	9871	3	US-09-479-122-24	Sequence 24, Appl	175	13.4	78.8	1549	3	US-09-668-885A-8	Sequence 8, Appl
103	13.8	81.2	9871	3	US-09-484-997-24	Sequence 24, Appl	176	13.4	78.8	1615	3	US-09-820-002-1	Sequence 1, Appl
104	13.8	81.2	9871	3	US-09-481-355-24	Sequence 24, Appl	c 177	13.4	78.8	1615	3	US-09-902-540-8917	Sequence 8917, Ap
105	13.8	81.2	9871	3	US-09-481-282-24	Sequence 24, Appl	c 178	13.4	78.8	1783	3	US-09-510-738A-188	Sequence 188, App
106	13.8	81.2	9871	3	US-09-455-659A-24	Sequence 24, Appl	179	13.4	78.8	1783	3	US-09-861-966-188	Sequence 188, App
107	13.8	81.2	9871	3	US-09-484-996-24	Sequence 24, Appl	180	13.4	78.8	1783	3	US-09-919-048-188	Sequence 188, App
108	13.8	81.2	9871	3	US-09-479-123-24	Sequence 24, Appl	181	13.4	78.8	1783	3	US-09-949-016-1311	Sequence 1311, Ap
109	13.8	81.2	9871	3	US-09-484-317A-24	Sequence 24, Appl	182	13.4	78.8	1783	3	US-10-102-283-188	Sequence 188, App
110	13.8	81.2	9871	3	US-09-276-820A-24	Sequence 24, Appl	c 183	13.4	78.8	1814	2	US-08-720-899-5	Sequence 5, Appl
111	13.8	81.2	9871	5	US-09-760-897A-24	Sequence 24, Appl	c 184	13.4	78.8	1814	2	US-08-459-610-5	Sequence 5, Appl
112	13.8	81.2	10060	3	US-09-479-122-25	Sequence 25, Appl	c 185	13.4	78.8	1814	2	US-08-343-804-5	Sequence 5, Appl
113	13.8	81.2	10060	3	US-09-484-997-25	Sequence 25, Appl	c 186	13.4	78.8	1814	2	US-08-687-399-5	Sequence 5, Appl
114	13.8	81.2	10060	3	US-09-481-355-25	Sequence 25, Appl	c 187	13.4	78.8	1814	2	US-08-600-908A-5	Sequence 5, Appl
115	13.8	81.2	10060	3	US-09-481-282-25	Sequence 25, Appl	c 188	13.4	78.8	1814	3	US-08-683-838A-5	Sequence 5, Appl
116	13.8	81.2	10060	3	US-09-455-659A-25	Sequence 25, Appl	c 189	13.4	78.8	1814	3	US-09-182-859-5	Sequence 5, Appl
117	13.8	81.2	10060	3	US-09-484-996-25	Sequence 25, Appl	c 190	13.4	78.8	1814	3	US-09-672-459-5	Sequence 5, Appl
118	13.8	81.2	10060	3	US-09-479-123-25	Sequence 25, Appl	c 191	13.4	78.8	1814	3	US-09-636-252A-5	Sequence 5, Appl
119	13.8	81.2	10060	3	US-09-484-317A-25	Sequence 25, Appl	c 192	13.4	78.8	1814	3	US-10-186-042-5	Sequence 5, Appl
120	13.8	81.2	10060	3	US-09-276-820A-25	Sequence 25, Appl	c 193	13.4	78.8	2066	3	US-10-104-047-56	Sequence 56, Appl
121	13.8	81.2	10060	5	US-09-760-897A-25	Sequence 25, Appl	194	13.4	78.8	2252	3	US-09-949-016-4519	Sequence 4519, Ap
122	13.8	81.2	10325	3	US-09-182-145-53	Sequence 53, Appl	195	13.4	78.8	2273	3	US-09-949-016-19	Sequence 19, Appl
123	13.8	81.2	14641	3	US-09-949-016-15834	Sequence 15834, A	196	13.4	78.8	2363	4	US-09-742-703-3	Sequence 3, Appl
124	13.8	81.2	17896	3	US-09-949-016-11896	Sequence 11896, A	c 197	13.4	78.8	2363	4	US-10-099-322-39	Sequence 39, Appl
125	13.8	81.2	18343	3	US-09-949-016-16413	Sequence 16413, A	198	13.4	78.8	2363	4	US-09-880-107-3690	Sequence 3690, Ap
126	13.8	81.2	21384	3	US-09-949-016-16283	Sequence 16283, A	c 199	13.4	78.8	2363	4	US-10-044-564-39	Sequence 39, Appl
127	13.8	81.2	23946	3	US-09-949-016-13381	Sequence 13381, A	c 200	13.4	78.8	2712	3	US-09-902-540-454	Sequence 454, App
128	13.8	81.2	28058	3	US-09-902-540-1252	Sequence 1252, Ap	c 201	13.4	78.8	2821	3	US-09-691-763B-1	Sequence 1, Appl
129	13.8	81.2	28762	3	US-09-902-540-1232	Sequence 1232, A	c 202	13.4	78.8	3336	10	5457037-1	Patent No. 5457037
130	13.8	81.2	29899	3	US-09-902-540-1265	Sequence 1265, Ap	c 203	13.4	78.8	5609	3	US-09-313-677-14	Sequence 14, Appl
131	13.8	81.2	36181	3	US-08-311-731A-120	Sequence 120, App	c 204	13.4	78.8	5609	3	US-09-313-677-15	Sequence 15, Appl
132	13.8	81.2	54779	3	US-09-949-001-27	Sequence 27, Appl	c 205	13.4	78.8	6196	3	US-09-902-540-788	Sequence 788, App
133	13.8	81.2	54780	3	US-09-949-001-39	Sequence 39, Appl	c 206	13.4	78.8	7477	3	US-09-902-540-950	Sequence 950, App
134	13.8	81.2	74644	3	US-09-949-016-17556	Sequence 17556, A	c 207	13.4	78.8	7707	3	US-09-949-016-11811	Sequence 11811, A
135	13.8	81.2	100848	3	US-09-596-002-39	Sequence 39, Appl	c 208	13.4	78.8	7707	3	US-09-949-016-16888	Sequence 16888, A
136	13.8	81.2	142783	3	US-09-949-016-15127	Sequence 15127, A	c 209	13.4	78.8	7742	2	US-08-221-750A-1	Sequence 1, Appl
137	13.8	81.2	144322	3	US-09-949-016-15316	Sequence 15316, A	c 210	13.4	78.8	8296	3	US-10-037-182-35	Sequence 35, Appl
138	13.8	81.2	154746	3	US-08-827-688-8	Sequence 8, Appl	c 211	13.4	78.8	8905	3	US-09-949-016-11761	Sequence 11761, A
139	13.8	81.2	232547	3	US-09-949-016-16603	Sequence 16603, A	c 212	13.4	78.8	8907	3	US-09-949-016-16261	Sequence 16261, A
140	13.8	81.2	4403765	3	US-09-103-840A-2	Sequence 2, Appl	c 213	13.4	78.8	9413	3	US-08-827-688-6	Sequence 6, Appl
141	13.8	81.2	4411529	3	US-09-103-840A-1	Sequence 1, Appl	c 214	13.4	78.8	10809	3	US-10-037-417-5	Sequence 5, Appl
142	13.4	78.8	20	3	US-09-742-703-46	Sequence 46, Appl	c 215	13.4	78.8	11220	3	US-09-949-016-5061	Sequence 5061, Ap
143	13.4	78.8	277	3	US-09-513-999C-34983	Sequence 34983, A	c 216	13.4	78.8	11220	3	US-10-037-182-1	Sequence 1, Appl
144	13.4	78.8	342	3	US-09-902-540-7869	Sequence 7869, Ap	c 217	13.4	78.8	11350	3	US-09-902-540-1179	Sequence 1179, Ap
145	13.4	78.8	405	3	US-08-854-133-706	Sequence 706, App	c 218	13.4	78.8	18195	3	US-09-820-002-3	Sequence 3, Appl
146	13.4	78.8	452	3	US-09-270-767-10935	Sequence 10935, A	c 219	13.4	78.8	21784	3	US-09-949-016-13053	Sequence 13053, A
147	13.4	78.8	593	3	US-09-621-976-455	Sequence 455, App	c 220	13.4	78.8	30337	3	US-09-949-016-14319	Sequence 14319, A
148	13.4	78.8	600	5	US-09-974-300-5866	Sequence 5866, Ap	c 221	13.4	78.8	31391	3	US-09-949-016-14319	Sequence 14319, A
149	13.4	78.8	601	3	US-09-949-016-77002	Sequence 77002, A	c 222	13.4	78.8	35574	3	US-09-949-016-11843	Sequence 11843, A
150	13.4	78.8	601	3	US-09-949-016-77003	Sequence 77003, A	c 223	13.4	78.8	35574	3	US-09-949-016-14511	Sequence 14511, A
151	13.4	78.8	601	3	US-09-949-016-77004	Sequence 77004, A	c 224	13.4	78.8	36093	3	US-09-949-016-14664	Sequence 14664, A
152	13.4	78.8	601	3	US-09-949-016-91513	Sequence 91513, A	c 225	13.4	78.8	36093	3	US-09-949-016-14665	Sequence 14665, A
153	13.4	78.8	601	3	US-09-949-016-164527	Sequence 164527, A	c 226	13.4	78.8	38371	3	US-09-949-016-12061	Sequence 12061, A
154	13.4	78.8	601	3	US-09-949-016-198701	Sequence 198701, A	c 227	13.4	78.8	38371	3	US-09-949-016-12488	Sequence 12488, A
155	13.4	78.8	626	3	US-09-691-763B-4	Sequence 4, Appl	c 228	13.4	78.8	38371	3	US-09-949-016-15596	Sequence 15596, A
156	13.4	78.8	858	2	US-08-796-676-2	Sequence 2, Appl	c 229	13.4	78.8	38371	3	US-09-949-016-15597	Sequence 15597, A
157	13.4	78.8	858	3	US-09-213-391-2	Sequence 2, Appl	c 230	13.4	78.8	40342	3	US-09-949-016-11821	Sequence 11821, A
158	13.4	78.8	861	3	US-09-489-039A-7088	Sequence 7088, Ap	c 231	13.4	78.8	40342	3	US-09-949-016-13982	Sequence 13982, A
159	13.4	78.8	1152	2	US-08-221-750A-10	Sequence 10, Appl	c 232	13.4	78.8	43102	3	US-09-949-016-16355	Sequence 16355, A
160	13.4	78.8	1215	3	US-09-489-039A-7121	Sequence 7121, Ap	c 233	13.4	78.8	46492	3	US-09-949-016-12953	Sequence 12953, A
161	13.4	78.8	1221	3	US-09-489-039A-2832	Sequence 2832, Ap	c 234	13.4	78.8	46492	3	US-09-949-016-12954	Sequence 12954, A
162	13.4	78.8	1455	3	US-09-902-540-6464	Sequence 6464, Ap	c 235	13.4	78.8	58162	3	US-09-949-016-16289	Sequence 16289, A
163	13.4	78.8	1548	3	US-08-600-656-6	Sequence 6, Appl	c 236	13.4	78.8	58789	3	US-09-949-016-15922	Sequence 15922, A
164	13.4	78.8	1548	3	US-09-170-670-11	Sequence 11, Appl	c 237	13.4	78.8	64984	3	US-09-949-016-15254	Sequence 15254, A
165	13.4	78.8	1548	3	US-09-193-068-11	Sequence 11, Appl	c 238	13.4	78.8	67002	3	US-09-949-016-16803	Sequence 16803, A
166	-13.4	78.8	1548	3	US-09-183-412-11	Sequence 11, Appl	c 239	13.4	78.8	67911	3	US-09-949-016-16979	Sequence 16979, A
167	13.4	78.8	1548	3	US-09-354-191A-6	Sequence 6, Appl	c 240	13.4	78.8	84132	3	US-09-949-016-16241	Sequence 16241, A
168	13.4	78.8	1548	3	US-09-290-734-11	Sequence 11, Appl	c 241	13.4	78.8	100990	3	US-09-409-800B-2	Sequence 2, Appl
169	13.4	78.8	1548	3	US-09-537-168-7	Sequence 7, Appl	c 242	13.4	78.8	103712	3	US-09-949-016-13058	Sequence 13058, A

C 243	13.4	78.8	107980	3	US-09-949-016-14370	Sequence 14370, A	316	12.8	75.3	597	5	US-09-854-867-272	Sequence 272, App
C 244	13.4	78.8	134008	3	US-09-949-016-13841	Sequence 13841, A	C 317	12.8	75.3	601	3	US-09-949-016-25416	Sequence 25416, A
C 245	13.4	78.8	143644	3	US-09-949-016-15238	Sequence 15238, A	C 318	12.8	75.3	601	3	US-09-949-016-33605	Sequence 33605, A
C 246	13.4	78.8	144596	3	US-09-949-016-11749	Sequence 11749, A	C 319	12.8	75.3	601	3	US-09-949-016-47630	Sequence 47630, A
C 247	13.4	78.8	144596	3	US-09-949-016-13095	Sequence 13095, A	C 320	12.8	75.3	601	3	US-09-949-016-53841	Sequence 53841, A
C 248	13.4	78.8	144922	3	US-09-949-016-15890	Sequence 15890, A	C 321	12.8	75.3	601	3	US-09-949-016-79425	Sequence 79425, A
C 249	13.4	78.8	187848	3	US-09-949-016-12111	Sequence 12111, A	C 322	12.8	75.3	601	3	US-09-949-016-109427	Sequence 109427, A
C 250	13.4	78.8	767677	3	US-09-949-016-12147	Sequence 12147, A	C 323	12.8	75.3	601	3	US-09-949-016-117774	Sequence 117774, A
C 251	13.4	78.8	767677	3	US-09-949-016-17361	Sequence 17361, A	C 324	12.8	75.3	601	3	US-09-949-016-117775	Sequence 117775, A
C 252	13	76.5	426	3	US-09-513-999C-32141	Sequence 32141, A	C 325	12.8	75.3	601	3	US-09-949-016-137208	Sequence 137208, A
C 253	13	76.5	504	3	US-09-252-991A-4754	Sequence 4754, Ap	C 326	12.8	75.3	601	3	US-09-949-016-137953	Sequence 137953, A
C 254	13	76.5	513	3	US-09-252-991A-4807	Sequence 4807, Ap	C 327	12.8	75.3	601	3	US-09-949-016-137954	Sequence 137954, A
C 255	13	76.5	537	3	US-09-489-039A-2727	Sequence 2727, Ap	C 328	12.8	75.3	601	3	US-09-949-016-137955	Sequence 137955, A
C 256	13	76.5	601	3	US-09-949-016-29052	Sequence 29052, A	C 329	12.8	75.3	601	3	US-09-949-016-137956	Sequence 137956, A
C 257	13	76.5	601	3	US-09-949-016-29053	Sequence 29053, A	C 330	12.8	75.3	601	3	US-09-949-016-137957	Sequence 137957, A
C 258	13	76.5	601	3	US-09-949-016-60918	Sequence 60918, A	C 331	12.8	75.3	601	3	US-09-949-016-138235	Sequence 138235, A
C 259	13	76.5	601	3	US-09-949-016-66059	Sequence 66059, A	C 332	12.8	75.3	601	3	US-09-949-016-150613	Sequence 150613, A
C 260	13	76.5	601	3	US-09-949-016-68594	Sequence 68594, A	C 333	12.8	75.3	601	3	US-09-949-016-150614	Sequence 150614, A
C 261	13	76.5	601	3	US-09-949-016-68595	Sequence 68595, A	C 334	12.8	75.3	601	3	US-09-949-016-153652	Sequence 153652, A
C 262	13	76.5	601	3	US-09-949-016-198700	Sequence 198700, A	C 335	12.8	75.3	601	3	US-09-949-016-153653	Sequence 153653, A
C 263	13	76.5	1014	3	US-09-252-991A-4690	Sequence 4690, Ap	C 336	12.8	75.3	601	3	US-09-949-016-153654	Sequence 153654, A
C 264	13	76.5	1844	3	US-10-104-047-1613	Sequence 1613, Ap	C 337	12.8	75.3	601	3	US-09-949-016-153655	Sequence 153655, A
C 265	13	76.5	3517	3	US-09-221-017B-281	Sequence 281, App	C 338	12.8	75.3	601	3	US-09-949-016-176839	Sequence 176839, A
C 266	13	76.5	4626	3	US-09-566-921-4	Sequence 4, Appli	C 339	12.8	75.3	601	3	US-09-949-016-176840	Sequence 176840, A
C 267	13	76.5	11221	4	US-09-778-963B-3	Sequence 3, Appli	C 340	12.8	75.3	601	3	US-09-949-016-198303	Sequence 198303, A
C 268	13	76.5	13048	4	US-09-880-107-2364	Sequence 2364, Ap	C 341	12.8	75.3	601	3	US-09-949-016-198304	Sequence 198304, A
C 269	13	76.5	26076	3	US-09-949-002-831	Sequence 831, App	C 342	12.8	75.3	601	3	US-09-949-016-198416	Sequence 198416, A
C 270	13	76.5	42244	3	US-09-949-016-13508	Sequence 13508, A	C 343	12.8	75.3	601	3	US-09-949-016-200142	Sequence 200142, A
C 271	13	76.5	54945	3	US-09-967-669-10	Sequence 10, Appl	C 344	12.8	75.3	613	3	US-09-328-111-556	Sequence 556, App
C 272	13	76.5	65485	3	US-09-949-016-13757	Sequence 13757, A	C 345	12.8	75.3	627	3	US-09-252-991A-12021	Sequence 12021, A
C 273	13	76.5	65552	3	US-09-949-016-12303	Sequence 12303, A	C 346	12.8	75.3	630	3	US-09-602-777A-145	Sequence 145, App
C 274	13	76.5	87562	3	US-09-949-016-13685	Sequence 13685, A	C 347	12.8	75.3	639	3	US-09-533-559-5006	Sequence 5006, App
C 275	13	76.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli	C 348	12.8	75.3	677	3	US-09-902-540-3211	Sequence 3211, Ap
C 276	13	76.5	441529	3	US-09-103-840A-1	Sequence 1, Appli	C 349	12.8	75.3	702	3	US-09-252-991A-11963	Sequence 11963, A
C 277	12.8	75.3	27	3	US-09-012-097A-46	Sequence 46, Appl	C 350	12.8	75.3	732	3	US-09-902-540-3211	Sequence 3211, Ap
C 278	12.8	75.3	27	3	US-09-481-620A-64	Sequence 64, Appl	C 351	12.8	75.3	756	3	US-09-634-238-83	Sequence 83, Appl
C 279	12.8	75.3	27	3	US-09-781-800A-45	Sequence 45, Appl	C 352	12.8	75.3	793	3	US-09-270-767-5906	Sequence 5906, Ap
C 280	12.8	75.3	36	3	US-09-694-127-8	Sequence 8, Appli	C 353	12.8	75.3	794	3	US-09-902-540-6886	Sequence 6886, Ap
C 281	12.8	75.3	41	3	US-09-485-737B-87	Sequence 87, Appl	C 354	12.8	75.3	794	3	US-09-902-540-6886	Sequence 6886, Ap
C 282	12.8	75.3	41	3	US-10-071-485-87	Sequence 87, Appl	C 355	12.8	75.3	807	3	US-09-443-501A-5	Sequence 5, Appli
C 283	12.8	75.3	59	2	US-08-356-161-15	Sequence 61, Appl	C 356	12.8	75.3	848	3	US-09-560-367A-3	Sequence 3, Appli
C 284	12.8	75.3	59	2	US-08-718-904-61	Sequence 61, Appl	C 357	12.8	75.3	848	3	US-09-724-889A-5	Sequence 5, Appli
C 285	12.8	75.3	59	3	US-09-449-249-61	Sequence 61, Appl	C 358	12.8	75.3	848	3	US-09-724-889A-5	Sequence 5, Appli
C 286	12.8	75.3	59	3	PCT-US93-05702-15	Sequence 15, Appl	C 359	12.8	75.3	848	3	US-09-957-483-3	Sequence 3, Appli
C 287	12.8	75.3	59	7	PCT-US95-10873A-15	Sequence 15, Appl	C 360	12.8	75.3	848	4	US-09-252-991A-5795	Sequence 5795, Ap
C 288	12.8	75.3	210	3	US-09-513-999C-24853	Sequence 24853, A	C 361	12.8	75.3	879	3	US-09-605-703B-2887	Sequence 2887, Ap
C 289	12.8	75.3	252	3	US-10-081-817A-23	Sequence 23, Appl	C 362	12.8	75.3	882	4	US-09-799-451-308	Sequence 308, App
C 290	12.8	75.3	258	3	US-10-081-817A-4	Sequence 4, Appli	C 363	12.8	75.3	961	3	US-09-252-991A-9964	Sequence 9964, Ap
C 291	12.8	75.3	287	3	US-09-270-767-26367	Sequence 26367, A	C 364	12.8	75.3	984	3	US-09-974-300-500	Sequence 500, App
C 292	12.8	75.3	312	3	US-10-081-817A-3	Sequence 3, Appli	C 365	12.8	75.3	987	5	US-10-258-472-1	Sequence 1, Appli
C 293	12.8	75.3	330	3	US-09-513-999C-2981	Sequence 2981, Ap	C 366	12.8	75.3	1011	3	US-09-252-991A-16085	Sequence 16085, A
C 294	12.8	75.3	384	3	US-09-533-559-7780	Sequence 7780, Ap	C 367	12.8	75.3	1026	3	US-09-252-991A-8069	Sequence 8069, Ap
C 295	12.8	75.3	402	3	US-09-252-991A-585	Sequence 585, App	C 368	12.8	75.3	1029	3	US-09-902-540-3651	Sequence 3651, Ap
C 296	12.8	75.3	454	2	US-08-253-155A-63	Sequence 63, Appl	C 369	12.8	75.3	1059	3	US-09-724-797-77	Sequence 77, Appl
C 297	12.8	75.3	454	3	US-09-533-559-2166	Sequence 2166, Ap	C 370	12.8	75.3	1065	3	US-09-252-991A-10622	Sequence 10622, A
C 298	12.8	75.3	479	3	US-09-513-999C-2014	Sequence 2014, Ap	C 371	12.8	75.3	1086	3	US-09-252-991A-9836	Sequence 9836, Ap
C 299	12.8	75.3	495	3	US-09-540-236-1896	Sequence 1896, Ap	C 372	12.8	75.3	1122	3	US-09-724-797-39	Sequence 39, Appl
C 300	12.8	75.3	500	2	US-08-809-237-45	Sequence 45, Appl	C 373	12.8	75.3	1137	3	US-09-252-991A-10118	Sequence 10118, A
C 301	12.8	75.3	537	3	US-09-252-991A-1106	Sequence 1106, Ap	C 374	12.8	75.3	1140	3	US-09-252-991A-1032	Sequence 1032, Ap
C 302	12.8	75.3	564	3	US-09-252-991A-8013	Sequence 8013, A	C 375	12.8	75.3	1143	3	US-09-602-777A-147	Sequence 147, App
C 303	12.8	75.3	570	3	US-09-991-181-407	Sequence 407, App	C 376	12.8	75.3	1155	3	US-09-134-000C-2903	Sequence 2903, Ap
C 304	12.8	75.3	570	3	US-09-990-444-407	Sequence 407, App	C 377	12.8	75.3	1158	3	US-09-974-300-244	Sequence 244, App
C 305	12.8	75.3	570	3	US-09-997-333-407	Sequence 407, App	C 378	12.8	75.3	1228	5	US-09-252-991A-12084	Sequence 12084, A
C 306	12.8	75.3	570	3	US-09-992-598-407	Sequence 407, App	C 379	12.8	75.3	1338	3	US-09-252-991A-8146	Sequence 8146, Ap
C 307	12.8	75.3	570	4	US-09-989-735-407	Sequence 407, App	C 380	12.8	75.3	1386	3	US-09-902-540-9088	Sequence 9088, Ap
C 308	12.8	75.3	570	5	US-09-989-726-407	Sequence 407, App	C 381	12.8	75.3	1413	3	US-09-280-116-10	Sequence 10, Appl
C 309	12.8	75.3	570	5	US-09-997-514-407	Sequence 407, App	C 382	12.8	75.3	1553	3	US-09-023-655-934	Sequence 934, App
C 310	12.8	75.3	570	5	US-09-989-728-407	Sequence 407, App	C 383	12.8	75.3	1559	3	US-09-949-016-3763	Sequence 3763, Ap
C 311	12.8	75.3	570	5	US-08-997-349-407	Sequence 407, App	C 384	12.8	75.3	1563	3	US-08-631-097-4	Sequence 4, Appli
C 312	12.8	75.3	570	5	US-09-997-653-407	Sequence 407, App	C 385	12.8	75.3	1568	2	US-08-810-712-3	Sequence 3, Appli
C 313	12.8	75.3	570	5	US-09-989-293A-407	Sequence 407, App	C 386	12.8	75.3	1568	3	US-08-853-733B-1	Sequence 1, Appli
C 314	12.8	75.3	591	3	US-09-533-559-474	Sequence 474, App	C 387	12.8	75.3	1599	3	US-08-602-791-1	Sequence 1, Appli
C 315	12.8	75.3	597	3	US-09-573-080A-272	Sequence 272, App	C 388	12.8	75.3	1600	3		

C 389	12.8	75.3	1603	2	US-08-625-209A-1	Sequence 1, Appli	C 462	12.8	75.3	5353	3	US-09-902-540-672	Sequence 672, App
C 390	12.8	75.3	1603	3	US-08-675-885-6	Sequence 6, Appli	C 463	12.8	75.3	5543	3	US-09-949-016-15615	Sequence 15615, A
C 391	12.8	75.3	1605	4	US-10-001-254-13	Sequence 13, Appl	C 464	12.8	75.3	5687	2	US-08-380-403A-3	Sequence 3, Appli
C 392	12.8	75.3	1617	3	US-09-252-991A-5856	Sequence 5856, Ap	C 465	12.8	75.3	5687	2	US-08-895-628-3	Sequence 3, Appli
C 393	12.8	75.3	1620	3	US-09-252-991A-8213	Sequence 8213, Ap	C 466	12.8	75.3	5687	3	US-08-895-810D-3	Sequence 3, Appli
C 394	12.8	75.3	1634	3	US-09-087-134-13	Sequence 13, Appl	C 467	12.8	75.3	6221	3	US-09-949-016-12307	Sequence 12307, A
C 395	12.8	75.3	1638	3	US-09-252-991A-494	Sequence 494, App	C 468	12.8	75.3	6222	3	US-09-949-016-13783	Sequence 13783, A
C 396	12.8	75.3	1648	4	US-09-880-107-3317	Sequence 3317, Ap	C 469	12.8	75.3	6336	4	US-09-880-107-1537	Sequence 1537, Ap
C 397	12.8	75.3	1659	3	US-09-252-991A-563	Sequence 563, App	C 470	12.8	75.3	6531	3	US-09-949-016-12002	Sequence 12002, A
C 398	12.8	75.3	1695	3	US-09-252-991A-534	Sequence 534, App	C 471	12.8	75.3	6532	3	US-09-949-016-14395	Sequence 14395, A
C 399	12.8	75.3	1695	3	US-09-252-991A-7816	Sequence 7816, Ap	C 472	12.8	75.3	6896	3	US-09-573-080A-422	Sequence 422, App
C 400	12.8	75.3	1830	3	US-09-149-476-129	Sequence 129, App	C 473	12.8	75.3	6896	5	US-09-854-867-422	Sequence 422, App
C 401	12.8	75.3	1836	3	US-09-252-991A-1026	Sequence 1026, Ap	C 474	12.8	75.3	6898	3	US-09-902-540-862	Sequence 862, App
C 402	12.8	75.3	1845	3	US-09-252-991A-8187	Sequence 8187, Ap	C 475	12.8	75.3	7460	3	US-09-949-016-12375	Sequence 12375, A
C 403	12.8	75.3	1846	5	US-10-130-323A-11	Sequence 11, Appl	C 476	12.8	75.3	7460	3	US-09-949-016-14384	Sequence 14384, A
C 404	12.8	75.3	1847	3	US-08-675-885-A-1	Sequence 4, Appli	C 477	12.8	75.3	7700	3	US-09-949-016-15619	Sequence 15619, A
C 405	12.8	75.3	1861	3	US-09-743-237-2	Sequence 2, Appli	C 478	12.8	75.3	9052	3	US-09-949-016-13662	Sequence 13662, A
C 406	12.8	75.3	1861	3	US-10-876-841A-2	Sequence 2, Appli	C 479	12.8	75.3	10054	3	US-09-902-540-971	Sequence 971, App
C 407	12.8	75.3	1929	3	US-10-876-841A-2	Sequence 2, Appli	C 480	12.8	75.3	10301	3	US-09-902-540-985	Sequence 985, App
C 408	12.8	75.3	1947	4	US-10-094-749-788	Sequence 788, App	C 481	12.8	75.3	10917	4	US-10-195-963-1	Sequence 1, Appli
C 409	12.8	75.3	2006	4	US-10-094-749-573	Sequence 573, App	C 482	12.8	75.3	11988	3	US-09-949-016-11977	Sequence 11977, A
C 410	12.8	75.3	2052	3	US-09-252-991A-10050	Sequence 10050, A	C 483	12.8	75.3	11989	3	US-09-949-016-13676	Sequence 13676, A
C 411	12.8	75.3	2100	3	US-09-836-932-2	Sequence 2, Appli	C 484	12.8	75.3	12163	3	US-09-949-016-16030	Sequence 16030, A
C 412	12.8	75.3	2103	3	US-09-252-991A-10237	Sequence 10237, A	C 485	12.8	75.3	12164	3	US-09-949-016-12202	Sequence 12202, A
C 413	12.8	75.3	2152	3	US-09-023-655-157	Sequence 157, App	C 486	12.8	75.3	12293	3	US-09-949-016-12222	Sequence 12222, A
C 414	12.8	75.3	2157	5	US-10-130-323A-2	Sequence 2, Appli	C 487	12.8	75.3	12293	3	US-09-949-016-13521	Sequence 13521, A
C 415	12.8	75.3	2241	3	US-09-743-237-1	Sequence 1, Appli	C 488	12.8	75.3	12899	3	US-09-949-016-16149	Sequence 16149, A
C 416	12.8	75.3	2241	3	US-10-876-841A-1	Sequence 1, Appli	C 489	12.8	75.3	13248	3	US-09-949-016-12024	Sequence 12024, A
C 417	12.8	75.3	2266	3	US-09-270-767-10887	Sequence 10887, A	C 490	12.8	75.3	13248	3	US-09-949-016-14230	Sequence 14230, A
C 418	12.8	75.3	2305	4	US-10-094-749-795	Sequence 795, App	C 491	12.8	75.3	13805	3	US-09-902-540-1083	Sequence 1083, Ap
C 419	12.8	75.3	2325	3	US-10-094-749-795	Sequence 795, App	C 492	12.8	75.3	15447	3	US-09-902-540-1100	Sequence 1100, Ap
C 420	12.8	75.3	2328	3	US-09-252-991A-8052	Sequence 8052, Ap	C 493	12.8	75.3	16356	3	US-09-949-016-15505	Sequence 15505, A
C 421	12.8	75.3	2334	3	US-09-252-991A-11890	Sequence 11890, A	C 494	12.8	75.3	16356	3	US-09-902-540-2000	Sequence 2000, Ap
C 422	12.8	75.3	2358	4	US-10-094-749-209	Sequence 209, App	C 495	12.8	75.3	16844	3	US-09-902-540-1176	Sequence 1176, Ap
C 423	12.8	75.3	2391	3	US-09-252-991A-5812	Sequence 5812, Ap	C 496	12.8	75.3	18367	3	US-09-949-016-14065	Sequence 14065, A
C 424	12.8	75.3	2511	5	US-10-130-323A-1	Sequence 1, Appli	C 497	12.8	75.3	18864	3	US-09-949-016-12119	Sequence 12119, A
C 425	12.8	75.3	2520	2	US-08-450-351-1	Sequence 1, Appli	C 498	12.8	75.3	18864	3	US-09-949-016-15598	Sequence 15598, A
C 426	12.8	75.3	2520	2	US-08-450-351-1	Sequence 3, Appli	C 499	12.8	75.3	20373	3	US-09-949-016-12302	Sequence 12302, A
C 427	12.8	75.3	2635	4	US-10-094-749-1514	Sequence 1514, Ap	C 500	12.8	75.3	20374	3	US-09-949-016-17086	Sequence 17086, A
C 428	12.8	75.3	2652	3	US-10-104-047-178	Sequence 178, App	C 501	12.8	75.3	21590	3	US-09-949-016-13372	Sequence 13372, A
C 429	12.8	75.3	2733	3	US-09-252-991A-7524	Sequence 7524, Ap	C 502	12.8	75.3	21590	3	US-09-949-016-13373	Sequence 13373, A
C 430	12.8	75.3	3027	3	US-09-328-352-3277	Sequence 3277, Ap	C 503	12.8	75.3	22372	3	US-09-949-016-17459	Sequence 17459, A
C 431	12.8	75.3	3074	3	US-09-710-279-3770	Sequence 3770, Ap	C 504	12.8	75.3	23381	3	US-09-949-016-13962	Sequence 13962, A
C 432	12.8	75.3	3142	5	US-09-410-362F-1	Sequence 1, Appli	C 505	12.8	75.3	23533	3	US-09-949-016-15377	Sequence 15377, A
C 433	12.8	75.3	3147	2	US-09-027-337-1	Sequence 1, Appli	C 506	12.8	75.3	23704	3	US-09-949-016-12104	Sequence 12104, A
C 434	12.8	75.3	3147	3	US-09-644-600-1	Sequence 1, Appli	C 507	12.8	75.3	23704	3	US-09-949-016-15312	Sequence 15312, A
C 435	12.8	75.3	3147	3	US-09-644-600-18	Sequence 18, Appl	C 508	12.8	75.3	24393	3	US-09-949-016-15431	Sequence 15431, A
C 436	12.8	75.3	3147	3	US-09-654-600A-1	Sequence 1, Appli	C 509	12.8	75.3	24395	3	US-09-949-016-14758	Sequence 14758, A
C 437	12.8	75.3	3147	3	US-09-654-600A-18	Sequence 18, Appli	C 510	12.8	75.3	24417	2	US-08-846-762-1	Sequence 1, Appli
C 438	12.8	75.3	3147	5	US-09-421-213-1	Sequence 1, Appli	C 511	12.8	75.3	27490	3	US-09-902-540-1227	Sequence 1227, Ap
C 439	12.8	75.3	3147	5	US-09-421-213-18	Sequence 18, Appl	C 512	12.8	75.3	27578	3	US-09-949-016-12167	Sequence 12167, A
C 440	12.8	75.3	3195	3	US-09-252-991A-937	Sequence 937, App	C 513	12.8	75.3	27579	3	US-09-949-016-15005	Sequence 15005, A
C 441	12.8	75.3	3386	3	US-09-041-886-1	Sequence 1, Appli	C 514	12.8	75.3	28626	3	US-09-596-002-15	Sequence 15, Appl
C 442	12.8	75.3	3386	3	US-09-949-016-362	Sequence 362, App	C 515	12.8	75.3	31571	2	US-08-323-443B-1	Sequence 1, Appli
C 443	12.8	75.3	3387	3	US-09-949-016-3570	Sequence 3570, Ap	C 516	12.8	75.3	35961	3	US-08-311-731A-138	Sequence 138, App
C 444	12.8	75.3	3435	3	US-09-758-759-166	Sequence 166, App	C 517	12.8	75.3	36171	3	US-09-949-016-13876	Sequence 13876, A
C 445	12.8	75.3	3442	3	US-09-902-540-6740	Sequence 6740, Ap	C 518	12.8	75.3	36274	3	US-09-949-016-12389	Sequence 12389, A
C 446	12.8	75.3	3443	3	US-09-902-540-542	Sequence 542, App	C 519	12.8	75.3	36470	3	US-08-311-731A-123	Sequence 123, App
C 447	12.8	75.3	3641	3	US-09-949-016-3877	Sequence 3877, Ap	C 520	12.8	75.3	36907	3	US-09-949-002-749	Sequence 749, App
C 448	12.8	75.3	3751	3	US-09-902-540-586	Sequence 586, App	C 521	12.8	75.3	37068	3	US-09-949-016-12543	Sequence 12543, A
C 449	12.8	75.3	3808	2	US-08-916-917-3	Sequence 3, Appli	C 522	12.8	75.3	37068	3	US-09-949-016-17376	Sequence 17376, A
C 450	12.8	75.3	3808	2	US-08-972-631-3	Sequence 3, Appli	C 523	12.8	75.3	37254	3	US-09-949-016-15973	Sequence 15973, A
C 451	12.8	75.3	3808	2	US-08-972-629-3	Sequence 3, Appli	C 524	12.8	75.3	38494	3	US-08-311-731A-24	Sequence 24, Appl
C 452	12.8	75.3	3808	2	US-08-972-630-3	Sequence 3, Appli	C 525	12.8	75.3	39113	3	US-09-949-016-15634	Sequence 15634, A
C 453	12.8	75.3	3808	2	US-08-672-211-3	Sequence 3, Appli	C 526	12.8	75.3	39376	3	US-09-949-016-17536	Sequence 17536, A
C 454	12.8	75.3	3808	2	US-09-205-170-3	Sequence 3, Appli	C 527	12.8	75.3	40261	3	US-09-949-016-11773	Sequence 11773, A
C 455	12.8	75.3	3823	3	US-09-512-250C-1	Sequence 1, Appli	C 528	12.8	75.3	40265	3	US-09-949-016-16168	Sequence 16168, A
C 456	12.8	75.3	3936	3	US-09-902-540-7211	Sequence 7211, Ap	C 529	12.8	75.3	42610	3	US-09-949-016-13882	Sequence 13882, A
C 457	12.8	75.3	4444	3	US-09-879-833-3	Sequence 3, Appli	C 530	12.8	75.3	43435	3	US-09-949-016-12909	Sequence 12909, A
C 458	12.8	75.3	4485	2	US-09-023-655-1286	Sequence 1286, Ap	C 531	12.8	75.3	43507	3	US-09-949-016-13297	Sequence 13297, A
C 459	12.8	75.3	4488	2	US-08-441-430-1	Sequence 1, Appli	C 532	12.8	75.3	51336	3	US-09-949-016-16054	Sequence 16054, A
C 460	12.8	75.3	4567	3	US-09-949-016-20	Sequence 20, Appl	C 533	12.8	75.3	52667	3	US-09-949-016-12019	Sequence 12019, A
C 461	12.8	75.3	4589	3	US-09-949-016-4100	Sequence 4100, Ap	C 534	12.8	75.3	54576	3	US-09-949-016-15954	Sequence 15954, A

535	12.8	75.3	54576	3	US-09-949-016-15955	Sequence 15955, A	608	12.4	72.9	601	3	US-09-949-016-38956	Sequence 38956, A
536	12.8	75.3	59258	3	US-09-949-002-581	Sequence 581, App	c 609	12.4	72.9	601	3	US-09-949-016-45924	Sequence 45924, A
c 537	12.8	75.3	60304	3	US-09-949-016-12218	Sequence 12218, A	610	12.4	72.9	601	3	US-09-949-016-45787	Sequence 45787, A
c 538	12.8	75.3	60305	3	US-09-949-016-15791	Sequence 15791, A	611	12.4	72.9	601	3	US-09-949-016-46788	Sequence 46788, A
c 539	12.8	75.3	60788	3	US-09-949-016-16789	Sequence 16789, A	612	12.4	72.9	601	3	US-09-949-016-49988	Sequence 49988, A
c 540	12.8	75.3	61158	3	US-09-949-016-15041	Sequence 15041, A	613	12.4	72.9	601	3	US-09-949-016-56360	Sequence 56360, A
c 541	12.8	75.3	65902	3	US-09-949-002-609	Sequence 609, App	c 614	12.4	72.9	601	3	US-09-949-016-81949	Sequence 81949, A
c 542	12.8	75.3	75431	3	US-09-949-016-15122	Sequence 15122, A	c 615	12.4	72.9	601	3	US-09-949-016-81950	Sequence 81950, A
c 543	12.8	75.3	76562	3	US-09-949-016-17482	Sequence 17482, A	c 616	12.4	72.9	601	3	US-09-949-016-93749	Sequence 93749, A
c 544	12.8	75.3	88758	3	US-09-949-016-13502	Sequence 13502, A	c 617	12.4	72.9	601	3	US-09-949-016-95261	Sequence 95261, A
c 545	12.8	75.3	92344	3	US-09-949-016-15802	Sequence 15802, A	c 618	12.4	72.9	601	3	US-09-949-016-95439	Sequence 95439, A
c 546	12.8	75.3	96109	3	US-09-596-002-35	Sequence 35, Appl	c 619	12.4	72.9	601	3	US-09-949-016-95617	Sequence 95617, A
c 547	12.8	75.3	99629	3	US-09-596-002-37	Sequence 37, Appl	c 620	12.4	72.9	601	3	US-09-949-016-95795	Sequence 95795, A
c 548	12.8	75.3	109519	3	US-09-758-759-1	Sequence 1, Appli	621	12.4	72.9	601	3	US-09-949-016-116712	Sequence 116712, A
c 549	12.8	75.3	112239	3	US-09-949-016-11144	Sequence 11144, A	622	12.4	72.9	601	3	US-09-949-016-116713	Sequence 116713, A
c 550	12.8	75.3	118067	3	US-09-497-858A-32	Sequence 32, Appl	623	12.4	72.9	601	3	US-09-949-016-120373	Sequence 120373, A
c 551	12.8	75.3	140844	3	US-09-949-016-14199	Sequence 14199, A	624	12.4	72.9	601	3	US-09-949-016-121007	Sequence 121007, A
c 552	12.8	75.3	162914	3	US-09-949-016-15578	Sequence 15578, A	625	12.4	72.9	601	3	US-09-949-016-121714	Sequence 121714, A
c 553	12.8	75.3	201529	3	US-09-949-016-12740	Sequence 12740, A	626	12.4	72.9	601	3	US-09-949-016-127726	Sequence 127726, A
c 554	12.8	75.3	211049	3	US-09-949-016-15770	Sequence 15770, A	627	12.4	72.9	601	3	US-09-949-016-128063	Sequence 128063, A
c 555	12.8	75.3	222691	3	US-09-949-016-11762	Sequence 11762, A	628	12.4	72.9	601	3	US-09-949-016-128757	Sequence 128757, A
c 556	12.8	75.3	222897	3	US-09-949-016-15842	Sequence 15842, A	629	12.4	72.9	601	3	US-09-949-016-128758	Sequence 128758, A
c 557	12.8	75.3	229354	3	US-09-705-400-64	Sequence 64, Appl	630	12.4	72.9	601	3	US-09-949-016-132153	Sequence 132153, A
c 558	12.8	75.3	269223	3	US-09-596-002-41	Sequence 41, Appl	c 631	12.4	72.9	601	3	US-09-949-016-132330	Sequence 132330, A
c 559	12.8	75.3	455726	3	US-09-949-016-14157	Sequence 14157, A	c 632	12.4	72.9	601	3	US-09-949-016-132331	Sequence 132331, A
c 560	12.8	75.3	481115	3	US-09-949-016-11940	Sequence 11940, A	c 633	12.4	72.9	601	3	US-09-949-016-136529	Sequence 136529, A
c 561	12.8	75.3	536165	3	US-09-214-808-1	Sequence 1, Appli	c 634	12.4	72.9	601	3	US-09-949-016-139518	Sequence 139518, A
c 562	12.8	75.3	784019	3	US-09-949-016-14033	Sequence 14033, A	c 635	12.4	72.9	601	3	US-09-949-016-147923	Sequence 147923, A
c 563	12.8	75.3	828152	3	US-09-949-016-12777	Sequence 12777, A	636	12.4	72.9	601	3	US-09-949-016-147936	Sequence 147936, A
c 564	12.8	75.3	1830121	3	US-09-557-884-1	Sequence 1, Appli	637	12.4	72.9	601	3	US-09-949-016-147937	Sequence 147937, A
c 565	12.8	75.3	1830121	3	US-09-643-990A-1	Sequence 1, Appli	638	12.4	72.9	601	3	US-09-949-016-149943	Sequence 149943, A
c 566	12.8	75.3	1830121	3	US-10-158-865-1	Sequence 1, Appli	639	12.4	72.9	601	3	US-09-949-016-149944	Sequence 149944, A
c 567	12.4	72.9	30	3	US-09-628-730-9	Sequence 9, Appli	640	12.4	72.9	601	3	US-09-949-016-149945	Sequence 149945, A
c 568	12.4	72.9	50	3	US-10-131-827-6648	Sequence 6648, Ap	641	12.4	72.9	601	3	US-09-949-016-163479	Sequence 163479, A
c 569	12.4	72.9	50	3	US-10-131-827-7038	Sequence 7038, Ap	642	12.4	72.9	601	3	US-09-949-016-172835	Sequence 172835, A
c 570	12.4	72.9	50	5	US-10-131-831-6648	Sequence 6648, Ap	643	12.4	72.9	601	3	US-09-949-016-172836	Sequence 172836, A
c 571	12.4	72.9	50	5	US-10-131-831-7038	Sequence 7038, Ap	644	12.4	72.9	601	3	US-09-949-016-174719	Sequence 174719, A
c 572	12.4	72.9	202	3	US-09-313-294A-1045	Sequence 1045, Ap	c 645	12.4	72.9	601	3	US-09-949-016-177779	Sequence 177779, A
c 573	12.4	72.9	210	3	US-09-513-998C-24761	Sequence 24761, A	c 646	12.4	72.9	601	3	US-09-949-016-177780	Sequence 177780, A
c 574	12.4	72.9	317	3	US-09-621-976-2349	Sequence 2349, Ap	c 647	12.4	72.9	601	3	US-09-949-016-185446	Sequence 185446, A
c 575	12.4	72.9	322	4	US-09-880-107-62	Sequence 62, Appl	c 648	12.4	72.9	601	3	US-09-949-016-191554	Sequence 191554, A
c 576	12.4	72.9	339	3	US-09-621-976-2341	Sequence 2341, Ap	c 649	12.4	72.9	601	3	US-09-949-016-191732	Sequence 191732, A
c 577	12.4	72.9	342	3	US-09-252-991A-12903	Sequence 12903, A	c 650	12.4	72.9	601	3	US-09-949-016-191910	Sequence 191910, A
c 578	12.4	72.9	351	3	US-09-894-844-32	Sequence 32, Appl	c 651	12.4	72.9	601	3	US-09-949-016-192088	Sequence 192088, A
c 579	12.4	72.9	357	3	US-09-902-540-2032	Sequence 2032, Ap	c 652	12.4	72.9	601	3	US-09-949-016-192794	Sequence 192794, A
c 580	12.4	72.9	364	3	US-09-513-998C-2708	Sequence 2708, Ap	653	12.4	72.9	601	3	US-09-949-002-7673	Sequence 7673, Ap
c 581	12.4	72.9	379	3	US-09-949-002-7681	Sequence 7681, Ap	654	12.4	72.9	601	3	US-09-949-002-7674	Sequence 7674, Ap
c 582	12.4	72.9	386	3	US-09-949-002-7680	Sequence 7680, Ap	655	12.4	72.9	601	3	US-09-949-002-7677	Sequence 7677, Ap
c 583	12.4	72.9	395	5	US-09-974-300-3661	Sequence 3661, Ap	656	12.4	72.9	601	3	US-09-949-002-7678	Sequence 7678, Ap
c 584	12.4	72.9	399	3	US-09-270-767-15005	Sequence 15005, A	c 657	12.4	72.9	606	3	US-09-252-991A-6320	Sequence 6320, Ap
c 585	12.4	72.9	402	3	US-09-621-976-113	Sequence 113, App	c 658	12.4	72.9	639	3	US-09-252-991A-15920	Sequence 15920, Ap
c 586	12.4	72.9	414	3	US-09-621-976-17862	Sequence 17862, A	c 659	12.4	72.9	701	4	US-09-605-703B-1937	Sequence 1937, Ap
c 587	12.4	72.9	417	3	US-09-252-991A-8008	Sequence 8008, Ap	c 660	12.4	72.9	702	3	US-09-252-991A-12588	Sequence 12588, A
c 588	12.4	72.9	420	2	US-08-470-179-110	Sequence 110, App	c 661	12.4	72.9	714	3	US-09-252-991A-10604	Sequence 10604, A
c 589	12.4	72.9	436	3	US-09-060-756-161	Sequence 161, App	c 662	12.4	72.9	717	3	US-09-902-540-6449	Sequence 6449, Ap
c 590	12.4	72.9	436	3	US-09-670-314-161	Sequence 161, App	c 663	12.4	72.9	720	3	US-09-902-540-2481	Sequence 2481, Ap
c 591	12.4	72.9	436	4	US-10-259-678-161	Sequence 161, App	c 664	12.4	72.9	753	5	US-10-375-010-5	Sequence 5, Appli
c 592	12.4	72.9	444	3	US-10-125-258-9	Sequence 9, Appli	665	12.4	72.9	794	4	US-09-297-648-2857	Sequence 2857, Ap
c 593	12.4	72.9	472	3	US-09-621-976-478	Sequence 478, App	666	12.4	72.9	816	3	US-09-252-991A-10249	Sequence 10249, A
c 594	12.4	72.9	528	3	US-09-949-016-4872	Sequence 4872, Ap	c 667	12.4	72.9	840	3	US-09-252-991A-6402	Sequence 6402, Ap
c 595	12.4	72.9	531	3	US-09-902-540-3342	Sequence 3342, Ap	c 668	12.4	72.9	930	3	US-09-487-558B-289	Sequence 289, App
c 596	12.4	72.9	534	3	US-09-893-737-173	Sequence 173, App	c 669	12.4	72.9	933	3	US-08-987-743-1	Sequence 1, Appli
c 597	12.4	72.9	537	3	US-09-252-991A-5627	Sequence 5627, Ap	670	12.4	72.9	954	3	US-09-252-991A-14555	Sequence 14555, A
c 598	12.4	72.9	537	3	US-09-621-976-13837	Sequence 13837, A	c 671	12.4	72.9	955	3	US-09-620-312D-228	Sequence 228, App
c 599	12.4	72.9	543	3	US-09-902-540-3533	Sequence 3533, Ap	c 672	12.4	72.9	1044	3	US-09-248-796A-5137	Sequence 5137, Ap
c 600	12.4	72.9	553	3	US-09-949-002-9805	Sequence 9805, Ap	c 673	12.4	72.9	1097	3	US-09-771-161A-58	Sequence 58, Appl
c 601	12.4	72.9	553	3	US-09-533-559-1960	Sequence 1960, Ap	c 674	12.4	72.9	1125	4	US-09-252-991A-6150	Sequence 6150, Ap
c 602	12.4	72.9	601	3	US-09-820-002-14	Sequence 14, Appl	c 675	12.4	72.9	1140	4	US-10-056-790-7	Sequence 7, Appli
c 603	12.4	72.9	601	3	US-09-949-016-13741	Sequence 13741, A	c 676	12.4	72.9	1140	4	US-10-056-790-31	Sequence 31, Appli
c 604	12.4	72.9	601	3	US-09-949-016-13742	Sequence 13742, A	c 677	12.4	72.9	1152	3	US-09-252-991A-7652	Sequence 7652, Ap
c 605	12.4	72.9	601	3	US-09-949-016-20363	Sequence 20363, A	c 678	12.4	72.9	1173	3	US-09-221-017B-1000	Sequence 1000, Ap
c 606	12.4	72.9	601	3	US-09-949-016-20686	Sequence 20686, A	c 679	12.4	72.9	1191	3	US-09-252-991A-8126	Sequence 8126, Ap
c 607	12.4	72.9	601	3	US-09-949-016-37517	Sequence 37517, A	c 680	12.4	72.9	1197	3	US-09-543-681A-3467	Sequence 3467, Ap

c 681	12.4	72.9	1215	3	US-09-894-844-74	Sequence 74, Appl	754	12.4	72.9	3989	4	US-10-094-749-978	Sequence 978, App
c 682	12.4	72.9	1247	3	US-09-740-035-1	Sequence 1, Appl	c 755	12.4	72.9	4050	3	US-09-902-540-604	Sequence 604, App
c 683	12.4	72.9	1273	4	US-10-000-489-31	Sequence 31, Appl	c 756	12.4	72.9	4117	3	US-09-484-970B-2	Sequence 2, Appl
c 684	12.4	72.9	1273	4	US-09-932-095B-31	Sequence 31, Appl	c 757	12.4	72.9	4355	3	US-09-799-451-941	Sequence 941, App
c 685	12.4	72.9	1273	5	US-10-000-986A-31	Sequence 31, Appl	c 758	12.4	72.9	4529	3	US-09-799-451-940	Sequence 940, App
c 686	12.4	72.9	1284	3	US-08-252-931A-15802	Sequence 15802, A	c 759	12.4	72.9	4646	3	US-09-799-451-939	Sequence 939, App
c 687	12.4	72.9	1290	2	US-08-448-744-5	Sequence 5, Appl	c 760	12.4	72.9	5165	3	US-09-902-540-729	Sequence 729, App
c 688	12.4	72.9	1302	3	US-09-902-540-8043	Sequence 5, Appl	c 761	12.4	72.9	5166	3	US-09-799-451-67	Sequence 67, Appl
c 689	12.4	72.9	1308	3	US-08-987-743-5	Sequence 8043, Ap	c 762	12.4	72.9	5420	3	US-09-902-540-598	Sequence 598, App
c 690	12.4	72.9	1350	4	US-09-605-703B-2145	Sequence 2145, Ap	c 763	12.4	72.9	6253	3	US-09-949-016-12023	Sequence 12023, A
c 691	12.4	72.9	1374	3	US-09-252-931A-13227	Sequence 13227, Ap	c 764	12.4	72.9	6253	3	US-09-949-016-16829	Sequence 175, App
c 692	12.4	72.9	1383	3	US-08-252-931A-7737	Sequence 7737, Ap	c 765	12.4	72.9	6327	5	US-10-114-270-175	Sequence 175, App
c 693	12.4	72.9	1386	3	US-09-252-931A-10462	Sequence 10462, A	c 766	12.4	72.9	6396	3	US-08-949-016-3344	Sequence 3344, Ap
c 694	12.4	72.9	1473	3	US-09-252-931A-9679	Sequence 9679, Ap	c 767	12.4	72.9	6432	4	US-10-154-419-3	Sequence 3, Appl
c 695	12.4	72.9	1473	3	US-09-252-931A-14687	Sequence 14687, A	c 768	12.4	72.9	6588	4	US-09-949-016-1076	Sequence 1076, Ap
c 696	12.4	72.9	1477	3	US-09-949-016-655	Sequence 655, App	c 769	12.4	72.9	6588	4	US-10-154-419-1	Sequence 1, Appl
c 697	12.4	72.9	1488	2	US-08-834-655-3	Sequence 3, Appl	c 770	12.4	72.9	6768	4	US-09-902-540-819	Sequence 819, App
c 698	12.4	72.9	1488	3	US-08-834-033A-3	Sequence 3, Appl	c 771	12.4	72.9	6806	3	US-09-566-921-42	Sequence 42, Appl
c 699	12.4	72.9	1488	3	US-09-363-574-3	Sequence 3, Appl	c 772	12.4	72.9	6856	3	US-09-962-665-2	Sequence 2, Appl
c 700	12.4	72.9	1488	3	US-09-363-526-3	Sequence 3, Appl	c 773	12.4	72.9	6972	3	US-09-962-665-2	Sequence 2, Appl
c 701	12.4	72.9	1488	3	US-09-330-235-19	Sequence 19, Appl	c 774	12.4	72.9	6972	3	US-09-962-677-2	Sequence 2, Appl
c 702	12.4	72.9	1491	3	US-09-328-352-2842	Sequence 15068, A	c 775	12.4	72.9	7538	3	US-09-902-540-894	Sequence 894, App
c 703	12.4	72.9	1509	3	US-09-328-352-2842	Sequence 2842, Ap	c 776	12.4	72.9	7538	3	US-08-451-777A-7	Sequence 7, Appl
c 704	12.4	72.9	1536	3	US-08-352-930-17	Sequence 17, Appl	c 777	12.4	72.9	7676	2	US-08-451-778A-7	Sequence 7, Appl
c 705	12.4	72.9	1536	3	US-09-573-080A-335	Sequence 335, App	c 778	12.4	72.9	7676	2	US-08-988-208-7	Sequence 7, Appl
c 706	12.4	72.9	1536	5	US-09-854-867-335	Sequence 335, App	c 779	12.4	72.9	7676	7	PCT-US95-06743-7	Sequence 7, Appl
c 707	12.4	72.9	1560	3	US-09-252-931A-8114	Sequence 8114, A	c 780	12.4	72.9	7702	3	US-09-023-655-1336	Sequence 1336, Ap
c 708	12.4	72.9	1671	3	US-09-252-931A-11030	Sequence 11030, A	c 781	12.4	72.9	7702	3	US-09-743-492A-3	Sequence 3, Appl
c 709	12.4	72.9	1695	3	US-09-902-540-8849	Sequence 8849, Ap	c 782	12.4	72.9	7702	3	US-09-719-272-3	Sequence 3, Appl
c 710	12.4	72.9	1709	5	US-08-974-300-5135	Sequence 5135, Ap	c 783	12.4	72.9	7702	3	US-09-743-492-3	Sequence 3, Appl
c 711	12.4	72.9	1749	5	US-09-949-016-3403	Sequence 3403, Ap	c 784	12.4	72.9	7716	3	US-09-949-016-16614	Sequence 16614, A
c 712	12.4	72.9	1788	3	US-09-540-236-1164	Sequence 1164, Ap	c 785	12.4	72.9	8095	4	US-09-880-107-2228	Sequence 2228, Ap
c 713	12.4	72.9	1851	3	US-09-970-516-3	Sequence 3, Appl	c 786	12.4	72.9	8542	3	US-09-949-016-17488	Sequence 17488, A
c 714	12.4	72.9	1857	3	US-09-991-181-173	Sequence 173, App	c 787	12.4	72.9	8920	3	US-09-949-016-15145	Sequence 15145, Ap
c 715	12.4	72.9	1866	3	US-08-990-444-173	Sequence 173, App	c 788	12.4	72.9	9112	3	US-09-949-016-4126	Sequence 942, App
c 716	12.4	72.9	1866	3	US-09-997-333-173	Sequence 173, App	c 789	12.4	72.9	9993	3	US-09-902-540-912	Sequence 942, App
c 717	12.4	72.9	1866	3	US-09-997-333-173	Sequence 173, App	c 790	12.4	72.9	10233	3	US-09-949-016-16032	Sequence 16032, A
c 718	12.4	72.9	1866	3	US-09-992-598-173	Sequence 173, App	c 791	12.4	72.9	10545	3	US-09-526-133A-14	Sequence 14, Appl
c 719	12.4	72.9	1866	4	US-09-989-735-173	Sequence 173, App	c 792	12.4	72.9	10573	3	US-09-949-002-679	Sequence 3, Appl
c 720	12.4	72.9	1866	5	US-09-989-735-173	Sequence 173, App	c 793	12.4	72.9	10696	3	US-09-949-002-736	Sequence 736, App
c 721	12.4	72.9	1866	5	US-09-997-514-173	Sequence 173, App	c 794	12.4	72.9	11290	3	US-09-949-016-15327	Sequence 15327, A
c 722	12.4	72.9	1866	5	US-09-989-728-173	Sequence 173, App	c 795	12.4	72.9	12097	3	US-09-949-016-15162	Sequence 15162, A
c 723	12.4	72.9	1866	5	US-09-997-349-173	Sequence 173, App	c 796	12.4	72.9	12412	2	US-08-350-878-18	Sequence 18, Appl
c 724	12.4	72.9	1866	5	US-09-997-653-173	Sequence 173, App	c 797	12.4	72.9	12412	2	US-09-738-884-3	Sequence 3, Appl
c 725	12.4	72.9	1866	5	US-09-989-735-173	Sequence 173, App	c 798	12.4	72.9	13953	3	US-10-036-961A-3	Sequence 3, Appl
c 726	12.4	72.9	1988	3	US-10-104-047-1091	Sequence 1091, Ap	c 799	12.4	72.9	14340	3	US-09-949-016-16972	Sequence 16972, A
c 727	12.4	72.9	2064	3	US-09-252-931A-9616	Sequence 9616, Ap	c 800	12.4	72.9	14340	3	US-09-902-540-1110	Sequence 1110, Ap
c 728	12.4	72.9	2140	3	US-09-902-540-448	Sequence 448, App	c 801	12.4	72.9	15209	3	US-09-817-180-3	Sequence 3, Appl
c 729	12.4	72.9	2380	3	US-09-817-676A-13	Sequence 13, Appl	c 802	12.4	72.9	15209	3	US-10-003-235-3	Sequence 3, Appl
c 730	12.4	72.9	2517	3	US-08-733-360A-4	Sequence 4, Appl	c 803	12.4	72.9	15297	3	US-10-660-763-3	Sequence 3, Appl
c 731	12.4	72.9	2517	3	US-08-987-743-16	Sequence 16, Appl	c 804	12.4	72.9	15297	5	US-09-949-016-13063	Sequence 13063, A
c 732	12.4	72.9	2517	3	US-08-916-935-4	Sequence 4, Appl	c 805	12.4	72.9	16360	3	US-09-902-540-1156	Sequence 1156, Ap
c 733	12.4	72.9	2517	4	US-09-880-107-3279	Sequence 3279, Ap	c 806	12.4	72.9	16387	3	US-09-949-016-17309	Sequence 17309, A
c 734	12.4	72.9	2571	3	US-09-252-931A-8190	Sequence 8190, Ap	c 807	12.4	72.9	16387	3	US-09-949-016-13806	Sequence 13806, A
c 735	12.4	72.9	2571	3	US-09-470-276-53	Sequence 53, Appl	c 808	12.4	72.9	17226	3	US-09-949-016-15625	Sequence 15625, A
c 736	12.4	72.9	2612	4	US-10-094-749-187	Sequence 187, App	c 809	12.4	72.9	18326	3	US-09-902-540-1169	Sequence 1169, Ap
c 737	12.4	72.9	2703	2	US-08-743-637B-6	Sequence 6, Appl	c 810	12.4	72.9	18326	3	US-09-902-540-1169	Sequence 1169, Ap
c 738	12.4	72.9	2703	3	US-08-526-840B-6	Sequence 6, Appl	c 811	12.4	72.9	18662	3	US-09-949-016-14655	Sequence 14655, A
c 739	12.4	72.9	2974	2	US-08-290-978A-4	Sequence 4, Appl	c 812	12.4	72.9	18947	3	US-09-949-016-15106	Sequence 15106, A
c 740	12.4	72.9	2974	2	US-08-780-869-4	Sequence 4, Appl	c 813	12.4	72.9	19826	3	US-09-949-016-16973	Sequence 16973, A
c 741	12.4	72.9	2997	3	US-09-252-931A-7831	Sequence 7831, Ap	c 814	12.4	72.9	19826	3	US-09-949-016-16729	Sequence 16729, A
c 742	12.4	72.9	3162	3	US-09-252-931A-9569	Sequence 9569, Ap	c 815	12.4	72.9	19826	3	US-09-902-540-1223	Sequence 1223, Ap
c 743	12.4	72.9	3209	3	US-09-270-767-13797	Sequence 13797, A	c 816	12.4	72.9	21535	3	US-09-949-016-12827	Sequence 12827, A
c 744	12.4	72.9	3366	3	US-09-949-016-5746	Sequence 5746, Ap	c 817	12.4	72.9	21535	3	US-09-949-016-12828	Sequence 12828, A
c 745	12.4	72.9	3370	7	US-08-060-822A-5	Sequence 5, Appl	c 818	12.4	72.9	21535	3	US-09-949-016-12829	Sequence 12829, A
c 746	12.4	72.9	3370	7	PCT-US94-05257-5	Sequence 5, Appl	c 819	12.4	72.9	21535	3	US-09-949-016-13366	Sequence 13366, A
c 747	12.4	72.9	3382	3	US-09-771-161A-59	Sequence 59, Appl	c 820	12.4	72.9	21536	3	US-09-949-016-13367	Sequence 13367, A
c 748	12.4	72.9	3467	3	US-09-743-492A-1	Sequence 1, Appl	c 821	12.4	72.9	21536	3	US-09-949-016-13368	Sequence 13368, A
c 749	12.4	72.9	3467	3	US-09-719-272-1	Sequence 1, Appl	c 822	12.4	72.9	21536	3	US-09-949-016-13369	Sequence 13369, A
c 750	12.4	72.9	3583	3	US-09-743-492-1	Sequence 1, Appl	c 823	12.4	72.9	21561	3	US-09-949-002-641	Sequence 641, App
c 751	12.4	72.9	3583	3	US-09-949-016-857	Sequence 857, App	c 824	12.4	72.9	21561	3	US-09-949-002-797	Sequence 797, App
c 752	12.4	72.9	3642	3	US-08-956-171E-328	Sequence 328, App	c 825	12.4	72.9	21561	3	US-09-949-002-797	Sequence 797, App
c 753	12.4	72.9	3642	3	US-08-781-986A-328	Sequence 328, App	c 826	12.4	72.9	21784	3	US-09-820-002-3	Sequence 3, Appl

c 973 12.4 72.9 203475 3 US-09-949-016-17229
c 974 12.4 72.9 205044 3 US-09-949-016-15851
c 975 12.4 72.9 205044 3 US-09-949-016-15852
c 976 12.4 72.9 205044 3 US-09-949-016-15853
c 977 12.4 72.9 219964 3 US-09-949-016-15086
c 978 12.4 72.9 223471 3 US-09-949-016-12387
c 979 12.4 72.9 223471 3 US-09-949-016-12724
c 980 12.4 72.9 223471 3 US-09-949-016-12725
c 981 12.4 72.9 223471 3 US-09-949-016-11842
c 982 12.4 72.9 223471 3 US-09-949-016-12706
c 983 12.4 72.9 275110 3 US-09-949-016-16070
c 984 12.4 72.9 283538 3 US-09-949-016-13506
c 985 12.4 72.9 300598 3 US-09-949-016-11868
c 986 12.4 72.9 302604 3 US-09-949-016-14588
c 987 12.4 72.9 302604 3 US-09-949-016-14589
c 988 12.4 72.9 304533 3 US-09-949-016-15371
c 989 12.4 72.9 304533 3 US-09-949-016-15372
c 990 12.4 72.9 308362 3 US-09-949-016-17119
c 991 12.4 72.9 319608 3 US-09-539-333D-1
c 992 12.4 72.9 319608 3 US-09-679-409-1
c 993 12.4 72.9 374159 3 US-09-949-016-15868
c 994 12.4 72.9 450395 3 US-09-949-016-15473
c 995 12.4 72.9 536165 3 US-09-214-808-1
c 996 12.2 71.8 17 3 US-09-866-108A-9670
c 997 12.2 71.8 25 3 US-09-866-108A-14562
c 998 12.2 71.8 25 3 US-09-866-108A-14563
c 999 12.2 71.8 25 3 US-09-866-108A-14564
1000 12.2 71.8 25 3 US-09-866-108A-14565

ALIGNMENTS

RESULT 1
US-09-949-016-29173
; Sequence 29173, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29173
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29173

Query Match 100.0%; Score 17; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 14 GCCCAGCGTTGGCCGAG 30

RESULT 2
US-09-949-016-126957
; Sequence 126957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

Sequence 17229, A
Sequence 15851, A
Sequence 15852, A
Sequence 15853, A
Sequence 15086, A
Sequence 12387, A
Sequence 12724, A
Sequence 12725, A
Sequence 11842, A
Sequence 12706, A
Sequence 16070, A
Sequence 13506, A
Sequence 11868, A
Sequence 14588, A
Sequence 14589, A
Sequence 15371, A
Sequence 15372, A
Sequence 17119, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 15868, A
Sequence 15473, A
Sequence 1, Appli
Sequence 9670, Ap
Sequence 14562, A
Sequence 14563, A
Sequence 14564, A
Sequence 14565, A

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126957
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126957

Query Match 100.0%; Score 17; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 14 GCCCAGCGTTGGCCGAG 30

RESULT 3
US-09-188-930-249
; Sequence 249, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; AND METHODS FOR THEIR USE
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-188-930-249

Query Match 100.0%; Score 17; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 990 GCCCAGCGTTGGCCGAG 1006

RESULT 4
US-09-312-283C-249
; Sequence 249, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; AND METHODS FOR THEIR USE

; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-249

Query Match 100.0%; Score 17; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGCGAG 17
|||
DB 990 GCCCAGCTTGCCGCGAG 1006

RESULT 5

US-09-495-050A-223
; Sequence 223, Application US/09495050A
; Patent No. 6492505

; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy

; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice

; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US

; CURRENT APPLICATION NUMBER: US/09/495,050A

; CURRENT FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/118,318

; PRIOR FILING DATE: February 1, 1999

; NUMBER OF SEQ ID NOS: 305

; SOFTWARE: PERL Program

; SEQ ID NO 223

; LENGTH: 1228

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. 6492505 2681738CT1

US-09-495-050A-223

Query Match 100.0%; Score 17; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGCGAG 17
|||
DB 541 GCCCAGCTTGCCGCGAG 557

RESULT 6

US-09-949-016-15355
; Sequence 3613, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3613
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3613

Query Match 100.0%; Score 17; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGCGAG 17
|||
DB 1117 GCCCAGCTTGCCGCGAG 1133

RESULT 7

US-09-949-016-572
; Sequence 572, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 572

; LENGTH: 2456

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-572

Query Match 100.0%; Score 17; DB 3; Length 2456;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCCGCGAG 17
|||
DB 1271 GCCCAGCTTGCCGCGAG 1287

RESULT 8

US-09-949-016-15355
; Sequence 15355, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15355

; LENGTH: 26086

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-15355

Query Match 100.0%; Score 17; DB 3; Length 26086;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
|||||

Db 22913 GCCCAGCGTTGCCGAG 22929

RESULT 9

US-09-949-016-12314
; Sequence 12314, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12314
; LENGTH: 26238
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12314

Query Match 100.0%; Score 17; DB 3; Length 26238;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
|||||

Db 23067 GCCCAGCGTTGCCGAG 23083

RESULT 10

US-09-252-991A-11321
; Sequence 11321, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11321
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11321

Query Match 90.6%; Score 15.4; DB 3; Length 1527;
Best Local Similarity 94.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
|||||

Db 25 GCCCAGCGTTGCCGAG 41

RESULT 11

US-08-996-083-2
; Sequence 2, Application US/08996083A
; Patent No. 6124095
; GENERAL INFORMATION:
; APPLICANT: Magna, Holly
; APPLICANT: Schaffer, Paul
; APPLICANT: Lawton, Michael
; APPLICANT: Yocum, Sue
; APPLICANT: Mitchell, Peter
; APPLICANT: Hutchinson, Nancy
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHOHYDROLASE-2
; FILE REFERENCE: PF-0420 US
; CURRENT APPLICATION NUMBER: US/08/996,083A
; CURRENT FILING DATE: 1997-12-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0.
; SEQ ID NO 2
; LENGTH: 4183
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte Clone No. 6124095: 1388013
; PUBLICATION INFORMATION:
US-08-996-083-2

Query Match 90.6%; Score 15.4; DB 3; Length 4183;
Best Local Similarity 94.1%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
|||||

Db 3254 GCCCAGCGTTGCCGAG 3270

RESULT 12

US-09-429-516-2
; Sequence 2, Application US/09429516
; Patent No. 6251389
; GENERAL INFORMATION:
; APPLICANT: Magna, Holly
; APPLICANT: Schaffer, Paul
; APPLICANT: Lawton, Michael
; APPLICANT: Yocum, Sue
; APPLICANT: Mitchell, Peter
; APPLICANT: Hutchinson, Nancy
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOS
; PHOSPHOHYDROLASE-2
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/429,516
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,083
; FILING DATE:

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0420 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SATPF1008
; CLONE: 1388013
;
US-09-429-516-2
;
; Query Match 90.6%; Score 15.4; DB 3; Length 4183;
; Best Local Similarity 94.1%; Pred. No. 1.4e+02;
; Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
Qy 1 GCCCAGCTTGCCGAG 17
Db 3254 GCCCAGCTTGCCCAAG 3270
;
RESULT 13
US-09-949-016-157184/c
; Sequence 157184, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157184
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-157184
;
; Query Match 84.7%; Score 14.4; DB 3; Length 601;
; Best Local Similarity 93.8%; Pred. No. 4.3e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
Qy 2 CCCACGCTTGCCGAG 17
Db 71 CCCACTTGCCGAG 56
;
RESULT 14
US-09-027-900-9/c
; Sequence 9, Application US/09027900
; Patent No. 6248551
; GENERAL INFORMATION:
; APPLICANT: De Reuse, Hilde
; APPLICANT: Skouloubri, Stephane
; APPLICANT: Labigne, Agnes
; TITLE OF INVENTION: HELICOBACTER ALIPHATIC AMIDASE Amie
; TITLE OF INVENTION: POLYPEPTIDES, AND DNA SEQUENCES ENCODING THOSE
; TITLE OF INVENTION: POLYPEPTIDES
;
US-09-027-900-8/c
; Sequence 8, Application US/09027900
; Patent No. 6248551
; GENERAL INFORMATION:
; APPLICANT: De Reuse, Hilde
; APPLICANT: Skouloubri, Stephane
; APPLICANT: Labigne, Agnes
; TITLE OF INVENTION: HELICOBACTER ALIPHATIC AMIDASE Amie
; TITLE OF INVENTION: POLYPEPTIDES, AND DNA SEQUENCES ENCODING THOSE
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,900
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,745
; FILING DATE: 28-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495-0165.00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1020 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1017
;
US-09-027-900-9
;
; Query Match 84.7%; Score 14.4; DB 3; Length 1020;
; Best Local Similarity 93.8%; Pred. No. 4.3e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
Qy 2 CCCACGCTTGCCGAG 17
Db 980 CCCACGCTTGCCGAG 965
;
RESULT 15
US-09-027-900-8/c
; Sequence 8, Application US/09027900
; Patent No. 6248551
; GENERAL INFORMATION:
; APPLICANT: De Reuse, Hilde
; APPLICANT: Skouloubri, Stephane
; APPLICANT: Labigne, Agnes
; TITLE OF INVENTION: HELICOBACTER ALIPHATIC AMIDASE Amie
; TITLE OF INVENTION: POLYPEPTIDES, AND DNA SEQUENCES ENCODING THOSE
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,900
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,745
; FILING DATE: 28-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495-0165.00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-027-900-8
;
; Query Match 84.7%; Score 14.4; DB 3; Length 1520;
; Best Local Similarity 93.8%; Pred. No. 4.3e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; Qy 2 CCCACGCTTGCCCGAG 17
; Db 1454 CCCACGCTTGCCCGAG 1439
;
; RESULT 16
; US-09-027-900-7/c
; Sequence 7, Application US/09027900
; Patent No. 6248551
; GENERAL INFORMATION:
; APPLICANT: De Reuse, Hilde
; APPLICANT: Skoulouris, Stephanie
; APPLICANT: Labigne, Agnes
; TITLE OF INVENTION: HELICOBACTER ALIPHATIC AMIDASE Amie
; TITLE OF INVENTION: POLYPEPTIDES, AND DNA SEQUENCES ENCODING THOSE
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,900
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,745
; FILING DATE: 28-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495-0165.00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1675 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-027-900-7
;
; Query Match 84.7%; Score 14.4; DB 3; Length 1675;
; Best Local Similarity 93.8%; Pred. No. 4.3e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; Qy 2 CCCACGCTTGCCCGAG 17
; Db 1534 CCCACGCTTGCCCGAG 1519
;
; RESULT 17
; US-09-949-016-16127/c
; Sequence 16127, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16127
; LENGTH: 42242
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16127
;
; Query Match 84.7%; Score 14.4; DB 3; Length 42242;
; Best Local Similarity 93.8%; Pred. No. 4.4e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; Qy 2 CCCACGCTTGCCCGAG 17
; Db 7247 CCCACTCTTGCCCGAG 7232
;
; RESULT 18
; US-09-949-016-13514
; Sequence 13514, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13514
; LENGTH: 47698
; TYPE: DNA
```

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (47698)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13514

Query Match      84.7%; Score 14.4; DB 3; Length 47698;
Best Local Similarity 93.8%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
Db 9925 CCCACGCTTGCTGAG 9940

RESULT 19
US-09-671-317-485/c
; Sequence 485, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62 US3 CIP
; CURRENT APPLICATION NUMBER: US/09/671,317
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm
; SEQ ID NO 485
; LENGTH: 49312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5466..7466
; OTHER INFORMATION: 5' regulatory region
; NAME/KEY: exon
; LOCATION: 7467..7725
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 20256..20355
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 36905..36975
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 45167..45248
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 45728..45965
; OTHER INFORMATION: exon 5
; NAME/KEY: misc feature
; LOCATION: 45966..49312
; OTHER INFORMATION: 3' regulatory region
; NAME/KEY: allele
; LOCATION: 7564
; OTHER INFORMATION: 10-286-289 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 7619
; OTHER INFORMATION: 10-286-345 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 7649
; OTHER INFORMATION: 10-286-345.mis complement

; OTHER INFORMATION: 10-286-375 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 17258
; OTHER INFORMATION: 12-425-57 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 21590
; OTHER INFORMATION: 12-421-135 : insertion of T
; NAME/KEY: allele
; LOCATION: 21595
; OTHER INFORMATION: 12-421-140 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 36971
; OTHER INFORMATION: 10-523-232 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 45214
; OTHER INFORMATION: 10-289-201 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 45741
; OTHER INFORMATION: 10-290-37 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 46029
; OTHER INFORMATION: 10-290-326 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 46032
; OTHER INFORMATION: 10-290-328 : deletion of G
; NAME/KEY: primer bind
; LOCATION: 7276..7294
; OTHER INFORMATION: 10-286.pu
; NAME/KEY: primer bind
; LOCATION: 7676..7694
; OTHER INFORMATION: 10-286.rp complement
; NAME/KEY: primer bind
; LOCATION: 16839..16856
; OTHER INFORMATION: 12-425.rp
; NAME/KEY: primer bind
; LOCATION: 17297..17314
; OTHER INFORMATION: 12-425.pu complement
; NAME/KEY: primer bind
; LOCATION: 21456..21474
; OTHER INFORMATION: 12-421.pu
; NAME/KEY: primer bind
; LOCATION: 21886..21906
; OTHER INFORMATION: 12-421.rp complement
; NAME/KEY: primer bind
; LOCATION: 36740..36758
; OTHER INFORMATION: 10-523.pu
; NAME/KEY: primer bind
; LOCATION: 36997..37015
; OTHER INFORMATION: 10-523.rp complement
; NAME/KEY: primer bind
; LOCATION: 45020..45037
; OTHER INFORMATION: 10-289.pu
; NAME/KEY: primer bind
; LOCATION: 45413..45432
; OTHER INFORMATION: 10-289.rp complement
; NAME/KEY: primer bind
; LOCATION: 45705..45724
; OTHER INFORMATION: 10-290.pu
; NAME/KEY: primer bind
; LOCATION: 46104..46123
; OTHER INFORMATION: 10-290.rp complement
; NAME/KEY: primer bind
; LOCATION: 7545..7563
; OTHER INFORMATION: 10-286-289.mis
; NAME/KEY: primer bind
; LOCATION: 7565..7583
; OTHER INFORMATION: 10-286-289.mis complement
; NAME/KEY: primer bind
; LOCATION: 7600..7618
; OTHER INFORMATION: 10-286-345.mis
; NAME/KEY: primer bind
; LOCATION: 7620..7638
; OTHER INFORMATION: 10-286-345.mis complement
```

```
; NAME/KEY: primer_bind
; LOCATION: 7630..7648
; OTHER INFORMATION: 10-286-375.mis
; NAME/KEY: primer_bind
; LOCATION: 7650..7668
; OTHER INFORMATION: 10-286-375.mis complement
; NAME/KEY: primer_bind
; LOCATION: 17239..17257
; OTHER INFORMATION: 12-425-57.mis
; NAME/KEY: primer_bind
; LOCATION: 17259..17277
; OTHER INFORMATION: 12-425-57.mis complement
; NAME/KEY: primer_bind
; LOCATION: 21576..21594
; OTHER INFORMATION: 12-421-140.mis
; NAME/KEY: primer_bind
; LOCATION: 21596..21614
; OTHER INFORMATION: 12-421-140.mis complement
; NAME/KEY: primer_bind
; LOCATION: 36952..36970
; OTHER INFORMATION: 10-523-232.mis
; NAME/KEY: primer_bind
; LOCATION: 36972..36990
; OTHER INFORMATION: 10-523-232.mis complement
; NAME/KEY: primer_bind
; LOCATION: 45195..45213
; OTHER INFORMATION: 10-289-201.mis
; NAME/KEY: primer_bind
; LOCATION: 45215..45233
; OTHER INFORMATION: 10-289-201.mis complement
; NAME/KEY: primer_bind
; LOCATION: 45722..45740
; OTHER INFORMATION: 10-290-37.mis
; NAME/KEY: primer_bind
; LOCATION: 45742..45760
; OTHER INFORMATION: 10-290-37.mis complement
; NAME/KEY: primer_bind
; LOCATION: 46010..46028
; OTHER INFORMATION: 10-290-326.mis
; NAME/KEY: primer_bind
; LOCATION: 46030..46048
; OTHER INFORMATION: 10-290-326.mis complement
; NAME/KEY: misc_binding
; LOCATION: 7552..7576
; OTHER INFORMATION: 10-286-289.probe
; NAME/KEY: misc_binding
; LOCATION: 7607..7631
; OTHER INFORMATION: 10-286-345.probe
; NAME/KEY: misc_binding
; LOCATION: 7637..7661
; OTHER INFORMATION: 10-286-375.probe
; NAME/KEY: misc_binding
; LOCATION: 17245..17270
; OTHER INFORMATION: 12-425-57.probe
; NAME/KEY: misc_binding
; LOCATION: 21583..21607
; OTHER INFORMATION: 12-421-140.probe
; NAME/KEY: misc_binding
; LOCATION: 36959..36983
; OTHER INFORMATION: 10-523-232.probe
; NAME/KEY: misc_binding
; LOCATION: 45202..45226
; OTHER INFORMATION: 10-289-201.probe
; NAME/KEY: misc_binding
; LOCATION: 45729..45753
; OTHER INFORMATION: 10-290-37.probe
; NAME/KEY: misc_binding
; LOCATION: 46017..46041
; OTHER INFORMATION: 10-290-326.probe
; US-09-671-317-485
```

Query Match 84.7%; Score 14.4; DB 3; Length 49312;
Best Local Similarity 93.8%; Pred. No. 4.4e+02;

```
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CCCACGCTTGGCCGAG 17
|||||
Db 12970 CCCACTCTTGGCCGAG 12955
```

RESULT 20

US-09-949-016-16803/c
; Sequence 16803, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16803
; LENGTH: 67002
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(67002)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16803

Query Match 84.7%; Score 14.4; DB 3; Length 67002;
Best Local Similarity 93.8%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 2 CCCACGCTTGGCCGAG 17
|||||
Db 41452 CCCAGGCTTGGCCGAG 41437
```

RESULT 21

US-09-332-522E-17/c
; Sequence 17, Application US/09332522E
; Patent No. 6781028
; GENERAL INFORMATION:
; APPLICANT: Costa, M.
; APPLICANT: Doberstein, S.
; APPLICANT: Elson, S.
; APPLICANT: Ferguson, K.
; APPLICANT: Homberger, S.
; TITLE OF INVENTION: ANIMAL MODELS AND METHODS FOR ANALYSIS OF LIPID METABOLISM AND S
; FILE REFERENCE: 7326-101, EX99-004
; CURRENT APPLICATION NUMBER: US/09/332,522E
; CURRENT FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-332-522E-17

Query Match 81.2%; Score 13.8; DB 3; Length 20;
Best Local Similarity 88.2%; Pred. No. 8.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;


```
QY      1  GCCCAGCTTGGCCGAG 17
      ||||| ||||| |||||
Db      18  GCCCAGCATGACCGAG 2

RESULT 22
US-09-422-978-2426
; Sequence 2426, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CP1
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2426
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-10951-434 : polymorphic base T or C
US-09-422-978-2426

Query Match      81.2%; Score 13.8; DB 3; Length 47;
Best Local Similarity 88.2%; Pred. No. 8.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCCCAGCTTGGCCGAG 17
      ||||| ||||| |||||
Db      27  GCCCAGCTTGTACAG 43

RESULT 23
US-09-313-294A-2571/c
; Sequence 2571, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2571
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700552904H1
US-09-313-294A-2571

Query Match      81.2%; Score 13.8; DB 3; Length 283;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCCCAGCTTGGCCGAG 17
      ||||| ||||| |||||
Db      18  GCCCAGCATGACCGAG 2

RESULT 24
US-09-313-294A-5109/c
; Sequence 5109, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5109
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700349603H1
; NAME/KEY: unsure
; LOCATION: 232, 236, 244, 250, 278, 280, 284
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5109

Query Match      81.2%; Score 13.8; DB 3; Length 293;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCCCAGCTTGGCCGAG 17
      ||||| ||||| |||||
Db      180  GCCCACCCTTGCCCAAG 164

RESULT 25
US-08-990-823-44
; Sequence 44, Application US/08990823D
; Patent No. 6228371
; GENERAL INFORMATION:
; APPLICANT: Nano, Francis
; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
; FILE REFERENCE: 49086
; CURRENT APPLICATION NUMBER: US/08/990,823D
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: US 96/10375
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,254
; EARLIER FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIN Ver. 2.0
; SEQ ID NO 44
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-08-990-823-44

Query Match      81.2%; Score 13.8; DB 3; Length 385;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCCCAGCTTGGCCGAG 17
      ||||| ||||| |||||
Db      11  GCCCGCGCTTGCCCTAG 27

RESULT 26
US-09-477-135A-44
; Sequence 44, Application US/09477135A
```

```
; Patent No. 6572865
; GENERAL INFORMATION:
; APPLICANT: Nano, Francis
; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
; TITLE OF INVENTION: immunostimulatory Peptides
; FILE REFERENCE: 52888
; CURRENT APPLICATION NUMBER: US/09/477,135A
; PRIOR FILING DATE: 2000-01-03
; PRIOR APPLICATION NUMBER: 08990823
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 96/10375
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: 60/000,254
; PRIOR FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-477-135A-44

Query Match      81.2%; Score 13.8; DB 3; Length 385;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      11 GCCCGCGCTTGCGCTAG 27

RESULT 27
US-10-233-942-3
; Sequence 3, Application US/10233942
; Patent No. 6977178
; GENERAL INFORMATION:
; APPLICANT: Hodge, Timothy et al
; TITLE OF INVENTION: System for Automated Transgenic Screening
; FILE REFERENCE: 023131.41500
; CURRENT APPLICATION NUMBER: US/10/233,942
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 60/230,371
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 600
; TYPE: DNA
; ORGANISM: streptomyces alboniger
US-10-233-942-3

Query Match      81.2%; Score 13.8; DB 4; Length 600;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      275 GCCCGCGCATGCGCGAG 291

RESULT 28
US-09-945-952A-3
; Sequence 3, Application US/09945952A
; Patent No. 7011943
; GENERAL INFORMATION:
; APPLICANT: Hodge, Timothy A.
; TITLE OF INVENTION: System for Automated Transgenic Screening
; FILE REFERENCE: 023131.41500
; CURRENT APPLICATION NUMBER: US/09/945,952A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/230,371
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 600
; TYPE: DNA
; ORGANISM: streptomyces alboniger
US-09-945-952A-3

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 600
; TYPE: DNA
; ORGANISM: streptomyces alboniger
US-09-945-952A-3

Query Match      81.2%; Score 13.8; DB 5; Length 600;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      275 GCCCGCGCATGCGCGAG 291

RESULT 29
US-09-949-016-121284
; Sequence 121284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121284
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-121284

Query Match      81.2%; Score 13.8; DB 3; Length 601;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      351 GCCCATGCTTGCCCAAG 367

RESULT 30
US-09-949-016-205926
; Sequence 205926, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205926
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-205926
```

US-09-949-016-205926

Query Match 81.2%; Score 13.8; DB 3; Length 601;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
||||| ||||| |||
DB 25 GCCCAGCTTGGCCGAG 41

RESULT 31
US-09-949-016-205927
; Sequence 205927, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205927
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-205927

Query Match 81.2%; Score 13.8; DB 3; Length 601;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
||||| ||||| |||
DB 52 GCCCAGCTTGGCCGAG 68

RESULT 32
US-09-902-540-5464/c
; Sequence 5464, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5464
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5464

Query Match 81.2%; Score 13.8; DB 3; Length 795;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
||||| ||||| |||

Db 712 GCCCAGCTTGGCCGAG 696

RESULT 33
US-09-252-991A-6261/c
; Sequence 6261, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6261
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6261

Query Match 81.2%; Score 13.8; DB 3; Length 999;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
||||| ||||| |||
DB 343 GCCCAGCTTGGCCGAG 327

RESULT 34
US-09-252-991A-547/c
; Sequence 547, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 547
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-547

Query Match 81.2%; Score 13.8; DB 3; Length 1176;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
||||| ||||| |||
DB 668 GCCCAGCTTGGCCGAG 652

RESULT 35
US-09-540-236-266
; Sequence 266, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 266
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-266

Query Match 81.2%; Score 13.8; DB 3; Length 1182;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
||||| ||||| |||||
Db 1080 GCCCAAGCTTAGCCGAG 1096

RESULT 36
US-08-095-726-13
; Sequence 13, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huei-Che B
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-095-726-13

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
||||| ||||| |||||
Db 209 GCCCAGCGCTGGCCGG 225

RESULT 37
US-08-095-726-15
; Sequence 15, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huei-Che B
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-095-726-15

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
||||| ||||| |||||
Db 209 GCCCAGCGCTGGCCGG 225

RESULT 37
US-08-095-726-15
; Sequence 15, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huei-Che B
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-095-726-15

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
||||| ||||| |||||
Db 209 GCCCAGCGCTGGCCGG 225

RESULT 38
US-08-096-623A-13
; Sequence 13, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huei-Che B
; TITLE OF INVENTION: Glycosylation of Zeaxanthin and
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-096-623A-13

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
||||| ||||| |||||
Db 209 GCCCAGCGCTGGCCGG 225

RESULT 38
US-08-096-623A-13
; Sequence 13, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huei-Che B
; TITLE OF INVENTION: Glycosylation of Zeaxanthin and
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-096-623A-13

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTGGCCGAG 17
||||| ||||| |||||
Db 209 GCCCAGCGCTGGCCGG 225

ADDRESSEE: Welsh & Katz, Ltd.
STREET: 120 S. Riverside Plaza, 22nd Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,623A
FILING DATE: 22-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/805,061
FILING DATE: 09-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,921
FILING DATE: 28-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/562,674
FILING DATE: 03-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/525,551
FILING DATE: 18-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/487,613
FILING DATE: 02-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: AMO-006.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 655-1500
TELEFAX: (312) 655-1501
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1235 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
POSITION IN GENOME:
MAP POSITION: -19 to 1216
UNITS: bp
US-08-096-623A-13

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAG 17
|||||
Db 209 GCCACGCTTGGCCGAG 225

RESULT 39

US-08-096-623A-15
Sequence 15, Application US/08096623A
Patent No. 5684238
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L.
APPLICANT: Brinkhaus, Friedhelm L.
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H.
APPLICANT: Yarger, James G.
APPLICANT: Yen, Hwei-Che B.
TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:

ADDRESSEE: Welsh & Katz, Ltd.
STREET: 120 S. Riverside Plaza, 22nd Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,623A
FILING DATE: 22-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/805,061
FILING DATE: 09-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,921
FILING DATE: 28-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/562,674
FILING DATE: 03-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/525,551
FILING DATE: 18-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/487,613
FILING DATE: 02-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: AMO-006.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 655-1500
TELEFAX: (312) 655-1501
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1235 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
POSITION IN GENOME:
MAP POSITION: -19 to 1216
UNITS: bp
US-08-096-623A-15

Query Match 81.2%; Score 13.8; DB 2; Length 1235;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCACGCTTGGCCGAG 17
|||||
Db 209 GCCACGCTTGGCCGAG 225

RESULT 40

US-09-252-991A-1019/c
Sequence 1019, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

```
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1019
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1019

Query Match      81.2%; Score 13.8; DB 3; Length 1251;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGCCCGAG 17
      ||||||| |||||
Db      225 GCCCAGCGCTGGCGGTG 209

RESULT 41
US-08-681-129-1/C
; Sequence 1, Application US/08681129
; Patent No. 5738854
; GENERAL INFORMATION:
; APPLICANT: Mettenleiter, Thomas Cristoph
; TITLE OF INVENTION: Pseudorabies virus vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,129
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,446
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 424
; APPLICATION NUMBER: EP 92.203.079.6
; FILING DATE: 06-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gortley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1578 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; STRAIN: Kaplan
US-08-681-129-1

Query Match      81.2%; Score 13.8; DB 2; Length 1578;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGCCCGAG 17
      ||||||| |||||
Db      227 GCCCAGCGCGGCCGAG 211

RESULT 42
US-09-252-991A-569
; Sequence 569, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 569
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-569

Query Match      81.2%; Score 13.8; DB 3; Length 1608;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGCCCGAG 17
      ||||||| |||||
Db      1334 GCCCAGCGTTGCCCGAG 1350

RESULT 43
US-09-869-588-32
; Sequence 32, Application US/09869588
; Patent No. 6790657
; GENERAL INFORMATION:
; APPLICANT: Arya
; TITLE OF INVENTION: Lentivirus Vector System
; FILE REFERENCE: 59316
; CURRENT APPLICATION NUMBER: US/09/869,588
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: PCT/US00/00390
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: 60/115,247
; PRIOR FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1616
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: IRSS and
; OTHER INFORMATION: Puromycin sequences
US-09-869-588-32

Query Match      81.2%; Score 13.8; DB 3; Length 1616;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGCCCGAG 17
      ||||||| |||||
Db      1244 GCCCGCGCATGCCCGAG 1260

RESULT 44
US-09-252-991A-1111
; Sequence 1111, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
```

; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1111
; LENGTH: 1653
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1111

Query Match 81.2%; Score 13.8; DB 3; Length 1653;
Best Local Similarity 88.2%; Pred. No. 8.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCG 17
|||||
DB 252 GCCCAGCGTTGCCGCG 268

RESULT 45
US-09-165-042-2
; Sequence 2, Application US/09165042
; Patent No. 6100077
; GENERAL INFORMATION:
; APPLICANT: Sturley, Stephen L.
; APPLICANT: Oelkers, Peter
; TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL
; FILE REFERENCE: 0575/56331
; CURRENT APPLICATION NUMBER: US/09/165,042
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Yeast
US-09-165-042-2

Query Match 81.2%; Score 13.8; DB 3; Length 1976;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCG 17
|||||
DB 208 GCCCAGCGTTGCCGCG 224

RESULT 46
US-09-252-991A-531/c
; Sequence 531, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 531
; LENGTH: 1983
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-531

Query Match 81.2%; Score 13.8; DB 3; Length 1983;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCG 17
|||||
DB 693 GCCCAGCGTTGCCGCG 677

RESULT 47
US-09-252-991A-6426/c
; Sequence 6426, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6426
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1174), (1177)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-6426

Query Match 81.2%; Score 13.8; DB 3; Length 2049;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGCCGCG 17
|||||
DB 1520 GCCCAGCGTTGCCGCG 1504

RESULT 48
US-08-359-117-1/c
; Sequence 1, Application US/08358117
; Patent No. 5608147
; GENERAL INFORMATION:
; APPLICANT: Kaphammer, Bryan J.
; TITLE OF INVENTION: ttda Gene Selectable Markers in Plants and the
; TITLE OF INVENTION: Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,117
; FILING DATE: 15-DEC-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1405.0030001

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2058 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 751..1611
US-08-358-117-1

Query Match 81.2%; Score 13.8; DB 2; Length 2058;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1573 GCTCAGCGTTGGCCGAG 1557

RESULT 49
US-08-470-588-1/c
Sequence 1, Application US/08470588
Patent No. 6100446
GENERAL INFORMATION:
APPLICANT: STREIBER, WOLFGANG R.
APPLICANT: TIMMIS, KENNETH N.
APPLICANT: ZENK, MEINHART H.
TITLE OF INVENTION: MICROORGANISMS AND PLASMIDS FOR
TITLE OF INVENTION: 2,4-DICHLOROPHENOXACETIC ACID (2,4-D) MONOOXYGENASE
TITLE OF INVENTION: FORMATION AND PROCESS FOR THE PRODUCTION OF THESE PLASMIDS
TITLE OF INVENTION: AND STRAINS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
STREET: 745 FIFTH AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10151
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,588
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: LAWRENCE, WILLIAM F.
REGISTRATION NUMBER: 28,029
REFERENCE/DOCKET NUMBER: 514413-3526
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2058 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-470-588-1

Query Match 81.2%; Score 13.8; DB 3; Length 2058;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17

Db 1573 GCTCAGCGTTGGCCGAG 1557

RESULT 50
US-09-976-594-315/c
Sequence 315, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Buchbinder, Jenny
APPLICANT: Furness, Michael
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 315
LENGTH: 2169
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: incyte ID No. 6673549 223407.4
NAME/KEY: unsure
LOCATION: 1034, 1043
OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-315

Query Match 81.2%; Score 13.8; DB 3; Length 2169;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 268 GCCCAGCGTTGGCCGGG 252

RESULT 51
US-09-613-444-1
Sequence 1, Application US/09613444
Patent No. 6444427
GENERAL INFORMATION:
APPLICANT: Ludwig, Erwin H.
APPLICANT: Farese, Robert V.
APPLICANT: Innerarity, Thomas L.
APPLICANT: Cases, Sylvaine
TITLE OF INVENTION: Polymorphisms in a Diacylglycerol
TITLE OF INVENTION: Acyltransferase Gene, and Methods of Use Thereof.
FILE REFERENCE: 6510191US1
CURRENT APPLICATION NUMBER: US/09/613,444
CURRENT FILING DATE: 2000-07-11
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2174
TYPE: DNA
ORGANISM: Homo sapiens
US-09-613-444-1

Query Match 81.2%; Score 13.8; DB 3; Length 2174;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1684 GCCCAGCGTTGGCTGCG 1700

RESULT 52
US-10-142-516-1

```

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1036
; LENGTH: 2847
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1036

Query Match 81.2%; Score 13.8; DB 3; Length 2847;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0

Qy 1 GCCCAGCGCTTGCCGAG 17
Db 2660 GCCCAGCGCTGGCGTG 2676

RESULT 55
US-08-162-809-1
; Sequence 1, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; APPLICANT: Sajjadi, Fereydoun G.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,809
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(3..419, 421..2858)
US-08-162-809-1

Query Match 81.2%; Score 13.8; DB 2; Length 3133;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0

Qy 1 GCCCAGCGCTTGCCGAG 17
Db 2507 GGCCTCGCTTGCCGAG 2523

```

```
RESULT 56
US-09-902-540-4713/c
; Sequence 4713, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barty S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,893
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4713
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; US-09-902-540-4713

Query Match      81.2%; Score 13.8; DB 3; Length 3453;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCGGAG 17
Db 2623 GCCCAGCGCTGGCGGAG 2607
|||||

RESULT 57
US-10-094-749-957
; Sequence 957, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 957
; LENGTH: 3636
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-094-749-957

Query Match      81.2%; Score 13.8; DB 4; Length 3636;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

US-09-902-254/c
; Sequence 254, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 254
; LENGTH: 3814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1081605.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 3792, 3796-3801
; OTHER INFORMATION: a, t, c, g, or other
; US-09-919-039-254

Query Match      81.2%; Score 13.8; DB 3; Length 3814;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCGGAG 17
Db 598 GCCCAGCGCTGGCGGAG 582
|||||

RESULT 59
US-09-252-991A-930/c
; Sequence 930, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 930
; LENGTH: 4563
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-930

Query Match      81.2%; Score 13.8; DB 3; Length 4563;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCGGAG 17
Db 3035 GCCCAGCGCTGGCGGTG 3019
|||||
```

```
RESULT 60
US-09-479-122-21
; Sequence 21, Application US/09479122
; Patent No. 6410266
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003C
; CURRENT APPLICATION NUMBER: US/09/479,122
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-479-122-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCGCGAG 17
        ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGAG 1517

RESULT 61
US-09-484-997-21
; Sequence 21, Application US/09484997
; Patent No. 6524818
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003U
; CURRENT APPLICATION NUMBER: US/09/484,997
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-484-997-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCGCGAG 17
        ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGAG 1517

RESULT 62
US-09-481-355-21
; Sequence 21, Application US/09481355
; Patent No. 6524824
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003F
; CURRENT APPLICATION NUMBER: US/09/481,355
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-481-355-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCGCGAG 17
        ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGAG 1517

RESULT 63
US-09-481-282-21
; Sequence 21, Application US/09481282
; Patent No. 6541221
; GENERAL INFORMATION:
```

```
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-484-997-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCGCGAG 17
        ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGAG 1517

RESULT 62
US-09-481-355-21
; Sequence 21, Application US/09481355
; Patent No. 6524824
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003F
; CURRENT APPLICATION NUMBER: US/09/481,355
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-481-355-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCGCGAG 17
        ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGAG 1517

RESULT 63
US-09-481-282-21
; Sequence 21, Application US/09481282
; Patent No. 6541221
; GENERAL INFORMATION:
```

FEATURE:	
NAME/KEY:	modified_base
NAME/_KEY_:	modified_base
LOCATION:	(890)
TITLE OF INVENTION:	ENDOGENOUS GENES
TITLE OF INVENTION:	COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
AFFILIANT:	RONDELLI, JEREMY

```
; FILE REFERENCE: 0221-0003B
; CURRENT APPLICATION NUMBER: US/09/479,123
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-123-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGCGAG 1517

RESULT 67
US-09-484-317A-21
; Sequence 21, Application US/09484317A
; Patent No. 6740503
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERFF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: ATX-007CP4DV11
; CURRENT APPLICATION NUMBER: US/09/484,317A
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
```

```
US-09-484-317A-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGCGAG 1517

RESULT 68
US-09-276-820A-21
; Sequence 21, Application US/09276820A
; Patent No. 6897066
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERFF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/276,820A
; CURRENT FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-276-820A-21

Query Match      81.2%; Score 13.8; DB 3; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCTTGCCGCGAG 17
      ||||| ||||| ||||| |||||
Db      1501 GCCCGCGCATGCCGCGAG 1517

RESULT 69
US-09-760-897A-21
; Sequence 21, Application US/09760897A
; Patent No. 7033782
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERFF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: ATX7CP4D15CNRCE
; CURRENT APPLICATION NUMBER: US/09/760,897A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/515,124
; PRIOR FILING DATE: 2000-02-27
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
```

```

; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-760-897A-21

Query Match      81.2%; Score 13.8; DB 5; Length 5382;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCCGAG 17
Db      1501 GCCCGCGCATGGCCGAG 1517

RESULT 70
US-08-487-283A-4
; Sequence 4, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Rollins, Scott
; APPLICANT: Rother, Russell P.
; APPLICANT: Springhorn, Jeremy P.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Thomas, Thomas C.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
; OF INFLAMMATORY DISEASES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seth A. Fidel
; STREET: 25 Science Park (Alexion)
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06511
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.4Mb storage
; COMPUTER: Macintosh Centris 610
; OPERATING SYSTEM: System 7
; SOFTWARE: WordPerfect 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,283A
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,208
; FILING DATE: 02-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Seth A. Fidel.
; REGISTRATION NUMBER: 38,449
; REFERENCE/DOCKET NUMBER: ALX-152.1 CIP

; PRIOR APPLICATION INFORMATION:
; TELEPHONE: (203)776-1790
; TELEFAX: (203)772-3655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8540 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Other nucleic acid
; DESCRIPTION: Apex-3P Eukaryotic
; DESCRIPTION: Expression Vector
US-08-487-283A-4

Query Match      81.2%; Score 13.8; DB 3; Length 8540;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCCGAG 17
Db      3112 GCCCGCGCATGGCCGAG 3128

RESULT 71
PCT-US96-05611A-12
; Sequence 12, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Lenardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Pelfrey, Clara M.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: Modified Myelin Protein Molecules
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
; COMPUTER: Macintosh Centris 610
; OPERATING SYSTEM: System 7
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05611A
; FILING DATE: 02-MAY-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,644
; FILING DATE: May 2, 1995
; APPLICATION NUMBER: 08/431,648
; FILING DATE: May 2, 1995
; APPLICATION NUMBER: 08/482,114
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8540 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
```



```
; TOPOLOGY: Circular
; MOLECULE TYPE: Other nucleic acid
; DESCRIPTION: Apex-3P Eukaryotic
; DESCRIPTION: Expression Vector
PCT-US96-05611A-12

Query Match      81.2%; Score 13.8; DB 7; Length 8540;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCCGAG 17
    ||||| ||| |||||
Db 3112 GCCCGCATGCCCGAG 3128

RESULT 72
US-09-479-122-22
; Sequence 22, Application US/09479122
; Patent No. 6410266
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003C
; CURRENT APPLICATION NUMBER: US/09/479,122
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-122-22

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCCGAG 17
    ||||| ||| |||||
Db 8958 GCCCGCATGCCCGAG 8974

RESULT 73
US-09-479-122-23
; Sequence 23, Application US/09479122
; Patent No. 6410266
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003C
```

```
; CURRENT APPLICATION NUMBER: US/09/479,122
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-122-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCCGAG 17
    ||||| ||| |||||
Db 8958 GCCCGCATGCCCGAG 8974

RESULT 74
US-09-479-122-28
; Sequence 28, Application US/09479122
; Patent No. 6410266
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003C
; CURRENT APPLICATION NUMBER: US/09/479,122
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-122-28
```

```
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-484-997-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCCGAG 17
      ||||| ||| |||||
Db      8958  GCCCGCATGCCGAG 8974

RESULT 75
US-09-484-997-22
; Sequence 22, Application US/09484997
; Patent No. 6524818
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003J
; CURRENT APPLICATION NUMBER: US/09/484,997
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-484-997-22

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCCGAG 17
      ||||| ||| |||||
Db      8958  GCCCGCATGCCGAG 8974

RESULT 76
US-09-484-997-23
; Sequence 23, Application US/09484997
; Patent No. 6524818
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003J
; CURRENT APPLICATION NUMBER: US/09/484,997
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
```

```
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-484-997-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCCGAG 17
      ||||| ||| |||||
Db      8958  GCCCGCATGCCGAG 8974

RESULT 77
US-09-484-997-28
; Sequence 28, Application US/09484997
; Patent No. 6524818
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003J
; CURRENT APPLICATION NUMBER: US/09/484,997
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-484-997-28

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

[illegible]

```
RESULT 81
US-09-481-282-22
; Sequence 22, Application US/09481282
; Patent No. 6541221
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/481,282
; CURRENT FILING DATE: 2000-01-11
; PRIOR FILING DATE: 1998-09-24
; PRIOR FILING DATE: 1997-09-26
; PRIOR FILING DATE: 1997-09-26
; PRIOR FILING DATE: 1999-03-08
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-481-282-22
Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGTTGGCGGAG 17
      ||||| ||||| ||||| |||||
Db      8958  GCCCGCGCATGGCGGAG 8974

RESULT 82
US-09-481-282-23
; Sequence 23, Application US/09481282
; Patent No. 6541221
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/481,282
; CURRENT FILING DATE: 2000-01-11
; PRIOR FILING DATE: 1998-09-24
; PRIOR FILING DATE: 1997-09-26
; PRIOR FILING DATE: 1997-09-26
; PRIOR FILING DATE: 1999-03-08
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-481-282-23
Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGTTGGCGGAG 17
      ||||| ||||| ||||| |||||
Db      8958  GCCCGCGCATGGCGGAG 8974

RESULT 83
US-09-481-282-28
; Sequence 28, Application US/09481282
; Patent No. 6541221
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/481,282
; CURRENT FILING DATE: 2000-01-11
; PRIOR FILING DATE: 1998-09-24
; PRIOR FILING DATE: 1997-09-26
; PRIOR FILING DATE: 1997-09-26
; PRIOR FILING DATE: 1999-03-08
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-481-282-28
Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGTTGGCGGAG 17
      ||||| ||||| ||||| |||||
Db      8958  GCCCGCGCATGGCGGAG 8974

RESULT 84
US-09-455-659A-22
; Sequence 22, Application US/09455659A
; Patent No. 6602686
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCCCAGCTTGCCGAG 17
      ||||| ||| |||||
Db      8958  GCCCGCATGCCGAG 8974

RESULT 85
US-09-455-659A-23
; Sequence 23, Application US/09455659A
; Patent No. 6602686
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
```

```
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-455-659A-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCCCAGCTTGCCGAG 17
      ||||| ||| |||||
Db      8958  GCCCGCATGCCGAG 8974

RESULT 86
US-09-455-659A-28
; Sequence 28, Application US/09455659A
; Patent No. 6602686
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003A
; CURRENT APPLICATION NUMBER: US/09/455,659A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
```

```

; FILE REFERENCE: 0221-0003H
; CURRENT APPLICATION NUMBER: US/09/484,996
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1998-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-996-22

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGCTGGCCGAG 17
      ||||| ||||| ||||| |||||
Db      8958  GCCCGCGCATGCCGAG 8974

RESULT 89
US-09-484-996-28
; Sequence 28, Application US/09484996
; Patent No. 6623958
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003H
; CURRENT APPLICATION NUMBER: US/09/484,996
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1998-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-996-28

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGCTGGCCGAG 17
      ||||| ||||| ||||| |||||
Db      8958  GCCCGCGCATGCCGAG 8974

RESULT 88
US-09-484-996-23
; Sequence 23, Application US/09484996
; Patent No. 6623958
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003H
; CURRENT APPLICATION NUMBER: US/09/484,996
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1998-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-996-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGCTGGCCGAG 17
      ||||| ||||| ||||| |||||
Db      8958  GCCCGCGCATGCCGAG 8974

RESULT 90
US-09-479-123-22
; Sequence 22, Application US/09479123
; Patent No. 6670185
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003B
; CURRENT APPLICATION NUMBER: US/09/479,123
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
```

```
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-123-22

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCCGAG 17
Db      8958  GCCCGCATGCGCGAG 8974

RESULT 91
US-09-479-123-23
; Sequence 23, Application US/09479123
; Patent No. 6670185
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003B
; CURRENT APPLICATION NUMBER: US/09/479,123
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-123-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCCGAG 17
Db      8958  GCCCGCATGCGCGAG 8974

RESULT 92
US-09-479-123-28
; Sequence 28, Application US/09479123
; Patent No. 6670185
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003B
; CURRENT APPLICATION NUMBER: US/09/479,123
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-479-123-28

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCTTGCCGAG 17
Db      8958  GCCCGCATGCGCGAG 8974

RESULT 93
US-09-484-317A-22
; Sequence 22, Application US/09484317A
; Patent No. 6740503
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERP, BRUCE
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: ATX-007CP4DV11
; CURRENT APPLICATION NUMBER: US/09/484,317A
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
```



```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-276-820A-22

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
    ||||| ||||| ||||| |||||
Db 8958 GCCCGCGCATGCCGAG 8974

RESULT 97
US-09-276-820A-23
; Sequence 23, Application US/09276820A
; Patent No. 6897066
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/276,820A
; CURRENT FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-276-820A-23

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
    ||||| ||||| ||||| |||||
Db 8958 GCCCGCGCATGCCGAG 8974

RESULT 98
US-09-276-820A-28
; Sequence 28, Application US/09276820A
; Patent No. 6897066
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003US
; CURRENT APPLICATION NUMBER: US/09/276,820A
; CURRENT FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-276-820A-28

Query Match      81.2%; Score 13.8; DB 3; Length 9737;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGGCCGAG 17
    ||||| ||||| ||||| |||||
Db 8958 GCCCGCGCATGCCGAG 8974

RESULT 99
US-09-760-897A-22
; Sequence 22, Application US/09760897A
; Patent No. 7033782
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: ATX7CP4D15CNRC
; CURRENT APPLICATION NUMBER: US/09/760,897A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/515,124
; PRIOR FILING DATE: 2000-02-27
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
```

```
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-760-897A-22

Query Match      81.2%; Score 13.8; DB 5; Length 9737;
Best Local Similarity 88.2%; Pred. NO. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCCGAG 17
      ||||| ||||| |||||
Db      8958 GCCCGCGCATGCCGCGAG 8974

RESULT 100
US-09-760-897A-23
; Sequence 23, Application US/09760897A
; Patent No. 7033782
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: ATX7CP4D15CNRC
; CURRENT APPLICATION NUMBER: US/09/760,897A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/515,124
; PRIOR FILING DATE: 2000-02-27
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-760-897A-23

Query Match      81.2%; Score 13.8; DB 5; Length 9737;
Best Local Similarity 88.2%; Pred. NO. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCCGAG 17
      ||||| ||||| |||||
Db      8958 GCCCGCGCATGCCGCGAG 8974
```

Search completed: June 10, 2006, 15:39:30
Job time : 73.1083 secs

,
.
.
.
.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:26:46 ; Search time 651.1 Seconds
(without alignments)
320.825 Million cell updates/sec

Title: US-10-600-816-31

Perfect score: 17

Sequence: 1 gccacgcttgccgag 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications NA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	17	100.0	17	8	US-10-600-816-31
2	17	100.0	42	8	US-10-600-816-30
3	17	100.0	497	6	US-10-066-543-1937
C 4	17	100.0	552	6	US-10-066-543-1811
C 5	17	100.0	620	3	US-09-969-034-2222
C 6	17	100.0	634	3	US-09-969-034-3393
C 7	17	100.0	642	7	US-10-125-968-701
8	17	100.0	1071	9	US-10-112-615-134
9	17	100.0	1212	3	US-09-866-050A-249
10	17	100.0	1212	6	US-10-152-661-249
11	17	100.0	1228	7	US-10-313-542-223
12	17	100.0	1400	13	US-11-060-756-2418
13	17	100.0	1400	13	US-11-060-756-6690
14	17	100.0	1460	10	US-10-936-626-64
15	17	100.0	1460	10	US-10-938-061-64
16	17	100.0	1619	7	US-10-224-289-5
17	17	100.0	1619	9	US-10-935-190-43
18	17	100.0	1788	10	US-10-505-486-196
19	17	100.0	2297	9	US-10-775-920-13
20	17	100.0	2302	7	US-10-224-289-3
21	17	100.0	2302	8	US-10-240-425-405
22	17	100.0	2302	10	US-10-775-920-9
23	17	100.0	2302	10	US-10-510-507-2
24	17	100.0	2305	6	US-10-775-920-12
25	17	100.0	2316	6	US-10-176-847-59
26	17	100.0	2316	15	US-11-080-991-59
27	17	100.0	2446	9	US-10-775-920-11
28	17	100.0	2456	6	US-10-225-567A-453
29	17	100.0	2456	7	US-10-269-909-63
30	17	100.0	2456	7	US-10-269-909-64
31	17	100.0	2456	7	US-10-295-027-619
32	17	100.0	2456	8	US-10-600-816-2
33	17	100.0	2456	9	US-10-600-816-18
34	17	100.0	2456	9	US-10-775-920-10
35	17	100.0	2456	10	US-10-936-626-40
36	17	100.0	2456	10	US-10-938-061-40
37	17	100.0	2456	16	US-11-169-041-32
38	17	100.0	2593	7	US-10-264-043-834
39	17	100.0	4239	6	US-10-198-846-10424
C 40	15.4	90.6	1020	8	US-10-335-977-1640
C 41	15.4	90.6	1050	8	US-10-335-977-1641
C 42	15.4	90.6	1173	10	US-10-956-157-41
C 43	15.4	90.6	1173	10	US-10-956-157-5276
C 44	15.4	90.6	2306	8	US-10-398-221-3404
C 45	15.4	90.6	4183	3	US-09-757-716-2
46	15.4	90.6	4222	9	US-10-723-860-7430
47	15.4	90.6	63761	8	US-10-322-281-557
48	15.4	90.6	68732	8	US-10-322-281-560
49	15.4	90.6	495269	8	US-10-398-221-8
C 50	15.4	90.6	2256646	8	US-10-470-565-1
C 51	15.4	90.6	2944528	16	US-11-045-004-1
C 52	15.4	90.6	3011208	8	US-10-398-221-2058
C 53	15	88.2	337	9	US-10-357-930-57611
C 54	15	88.2	404	9	US-10-425-115-123425
C 55	15	88.2	3303	9	US-10-739-930-4894
C 56	14.4	84.7	25	10	US-10-956-157-193814
C 57	14.4	84.7	201	10	US-10-995-561-64090
C 58	14.4	84.7	344	3	US-09-803-719-429
C 59	14.4	84.7	344	10	US-10-779-543-13699
C 60	14.4	84.7	570	9	US-10-357-930-54041
C 61	14.4	84.7	653	4	US-09-925-065A-93940
C 62	14.4	84.7	653	5	US-09-925-065A-93940
C 63	14.4	84.7	834	4	US-09-925-065A-63251
C 64	14.4	84.7	834	5	US-09-925-065A-63251
C 65	14.4	84.7	834	12	US-10-301-480-164489
C 66	14.4	84.7	834	12	US-10-301-480-777898
C 67	14.4	84.7	1889	7	US-10-101-510-469
C 68	14.4	84.7	13608	3	US-09-861-846-3
C 69	14.4	84.7	13608	8	US-10-250-463-3
C 70	14.4	84.7	49312	8	US-10-294-934-485
C 71	14.4	84.7	49312	16	US-11-204-311-485
C 72	14.4	84.7	326002	10	US-10-461-862-56
73	14	82.4	25	8	US-10-713-956-525977
74	14	82.4	25	9	US-10-719-900-974704
C 75	14	82.4	201	10	US-10-995-561-63916
C 76	14	82.4	429	12	US-10-301-480-242988
C 77	14	82.4	429	12	US-10-301-480-856397
C 78	14	82.4	440	4	US-09-925-065A-148858
C 79	14	82.4	440	5	US-09-925-065A-148858
C 80	14	82.4	458	3	US-09-918-995-28686
C 81	14	82.4	465	9	US-10-674-124A-19328
C 82	14	82.4	518	5	US-09-925-065A-82285
C 83	14	82.4	518	5	US-09-925-065A-82285
C 84	14	82.4	518	12	US-10-301-480-183525
C 85	14	82.4	518	12	US-10-301-480-796934
C 86	14	82.4	614	4	US-09-925-065A-805681
C 87	14	82.4	614	5	US-09-925-065A-805681
C 88	14	82.4	683	4	US-09-925-065A-14
C 89	14	82.4	683	4	US-09-925-065A-15
90	14	82.4	683	5	US-09-925-065A-14

Sequence 196, App
Sequence 13, Appl
Sequence 3, Appl
Sequence 405, App
Sequence 9, Appl
Sequence 2, Appl
Sequence 12, Appl
Sequence 59, Appl
Sequence 11, Appl
Sequence 453, App
Sequence 63, Appl
Sequence 64, Appl
Sequence 619, App
Sequence 2, Appl
Sequence 10, Appl
Sequence 18, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 32, Appl
Sequence 834, App
Sequence 1640, Ap
Sequence 1641, Ap
Sequence 41, Appl
Sequence 5276, Ap
Sequence 3404, Ap
Sequence 2, Appl
Sequence 7430, Ap
Sequence 557, App
Sequence 560, App
Sequence 8, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 2058, Ap
Sequence 57611, A
Sequence 123425,
Sequence 4894, Ap
Sequence 193814,
Sequence 64090, A
Sequence 429, App
Sequence 13699, A
Sequence 54041, A
Sequence 93940,
Sequence 93940,
Sequence 63251, A
Sequence 63251, A
Sequence 164489,
Sequence 777898,
Sequence 469, App
Sequence 3, Appl
Sequence 3, Appl
Sequence 485, App
Sequence 485, App
Sequence 56, Appl
Sequence 525977,
Sequence 974704,
Sequence 63916, A
Sequence 242988,
Sequence 856397,
Sequence 148958,
Sequence 148858,
Sequence 28686, A
Sequence 19328, A
Sequence 82285, A
Sequence 82285, A
Sequence 183525,
Sequence 786934,
Sequence 805681,
Sequence 805681,
Sequence 14, Appl
Sequence 15, Appl
Sequence 14, Appl

91	14	82.4	683	5	US-09-925-065A-15	Sequence 101250, A	164	13.8	81.2	597	12	US-10-943-508A-15	Sequence 15, Appl
92	14	82.4	683	12	US-10-301-480-101250	Sequence 101250, A	c 165	13.8	81.2	599	9	US-10-425-115-59014	Sequence 59014, A
93	14	82.4	683	12	US-10-301-480-101251	Sequence 101251, A	166	13.8	81.2	600	3	US-09-945-592A-3	Sequence 3, Appl
94	14	82.4	683	12	US-10-301-480-714659	Sequence 714659, A	c 167	13.8	81.2	600	6	US-10-027-632-59372	Sequence 59372, A
95	14	82.4	683	12	US-10-301-480-714660	Sequence 714660, A	168	13.8	81.2	600	7	US-10-233-942-3	Sequence 3, Appl
96	14	82.4	911	12	US-10-301-480-598304	Sequence 598304, A	169	13.8	81.2	600	7	US-10-117-960-5	Sequence 5, Appl
97	14	82.4	911	12	US-10-301-480-1211713	Sequence 1211713, A	c 170	13.8	81.2	600	7	US-10-027-632-59372	Sequence 59372, A
98	14	82.4	1181	16	US-10-115-831-63	Sequence 63, Appl	171	13.8	81.2	600	8	US-10-437-963-102256	Sequence 102256, A
99	14	82.4	2475	16	US-11-096-568A-11881	Sequence 11881, A	172	13.8	81.2	600	10	US-10-972-079-5871	Sequence 5871, Ap
100	14	82.4	70513	10	US-10-995-561-13368	Sequence 13368, A	173	13.8	81.2	600	10	US-10-972-079-30611	Sequence 30611, A
101	14	82.4	561515	8	US-10-741-601-5682	Sequence 5682, Ap	174	13.8	81.2	600	10	US-10-972-079-30612	Sequence 30612, A
102	14	82.4	561515	9	US-10-741-600-17730	Sequence 17730, A	175	13.8	81.2	600	10	US-10-972-079-30613	Sequence 30613, A
103	14	82.4	2756646	8	US-10-470-565-1	Sequence 1, Appl	176	13.8	81.2	600	10	US-10-972-079-30614	Sequence 30614, A
104	14	82.4	27571748	8	US-10-297-465A-1	Sequence 1, Appl	177	13.8	81.2	600	10	US-10-972-079-30615	Sequence 30615, A
105	13.8	81.2	22	11	US-10-310-914A-131813	Sequence 131813, A	178	13.8	81.2	600	13	US-10-972-079-30616	Sequence 30616, A
106	13.8	81.2	25	11	US-10-719-956-624960	Sequence 624960, A	179	13.8	81.2	600	13	US-11-074-995-3	Sequence 3, Appl
107	13.8	81.2	25	9	US-10-719-900-11621	Sequence 11621, A	180	13.8	81.2	600	13	US-11-130-921-3	Sequence 3, Appl
108	13.8	81.2	47	7	US-10-349-143-2426	Sequence 2426, Ap	181	13.8	81.2	600	13	US-11-166-990-54	Sequence 54, Appl
109	13.8	81.2	201	16	US-11-124-367A-11233	Sequence 11233, A	c 182	13.8	81.2	600	16	US-11-136-527-5985	Sequence 5985, Ap
110	13.8	81.2	225	8	US-09-963-632-28401	Sequence 28401, A	183	13.8	81.2	600	16	US-11-170-693-54	Sequence 54, Appl
111	13.8	81.2	228	3	US-09-864-408A-6145	Sequence 6145, Ap	184	13.8	81.2	601	4	US-09-925-065A-819788	Sequence 819788, A
112	13.8	81.2	290	6	US-10-116-265-3	Sequence 3, Appl	185	13.8	81.2	601	5	US-09-925-065A-819788	Sequence 819788, A
113	13.8	81.2	290	3	US-09-732-627A-4248	Sequence 4248, Ap	186	13.8	81.2	615	4	US-09-925-065A-123716	Sequence 123716, A
114	13.8	81.2	316	9	US-10-435-115-157270	Sequence 157270, A	187	13.8	81.2	615	5	US-09-925-065A-123716	Sequence 123716, A
115	13.8	81.2	346	12	US-10-301-480-294557	Sequence 294557, A	188	13.8	81.2	621	12	US-10-943-508A-91	Sequence 91, Appl
116	13.8	81.2	346	12	US-10-301-480-907966	Sequence 907966, A	c 189	13.8	81.2	635	4	US-09-925-065A-482604	Sequence 482604, A
117	13.8	81.2	348	4	US-09-925-065A-206799	Sequence 206799, A	c 190	13.8	81.2	635	5	US-09-925-065A-482604	Sequence 482604, A
118	13.8	81.2	348	5	US-09-925-065A-206799	Sequence 206799, A	c 191	13.8	81.2	640	9	US-10-425-115-160447	Sequence 160447, A
119	13.8	81.2	350	3	US-09-867-701-1134	Sequence 1134, Ap	192	13.8	81.2	664	8	US-10-437-963-40592	Sequence 40592, A
120	13.8	81.2	363	3	US-09-864-408A-1437	Sequence 1437, Ap	193	13.8	81.2	677	7	US-10-032-189-13	Sequence 13, Appl
121	13.8	81.2	370	9	US-10-425-115-88184	Sequence 88184, A	194	13.8	81.2	679	8	US-10-767-701-10813	Sequence 10813, A
122	13.8	81.2	378	10	US-10-966-634-44	Sequence 966, A	c 195	13.8	81.2	690	6	US-10-027-632-164037	Sequence 164037, A
123	13.8	81.2	385	3	US-09-997-182-44	Sequence 44, Appl	c 196	13.8	81.2	690	6	US-10-027-632-164038	Sequence 164038, A
124	13.8	81.2	385	3	US-09-997-182-44	Sequence 44, Appl	c 197	13.8	81.2	690	6	US-10-027-632-164039	Sequence 164039, A
125	13.8	81.2	385	3	US-09-997-181-44	Sequence 44, Appl	c 198	13.8	81.2	690	7	US-10-027-632-164037	Sequence 164037, A
126	13.8	81.2	391	3	US-09-918-95-4818	Sequence 4818, Ap	c 199	13.8	81.2	690	7	US-10-027-632-164038	Sequence 164038, A
127	13.8	81.2	395	8	US-10-424-539-45393	Sequence 45393, A	c 200	13.8	81.2	735	4	US-09-925-065A-23101	Sequence 23101, A
128	13.8	81.2	411	4	US-09-925-065A-562676	Sequence 562676, A	c 201	13.8	81.2	735	4	US-09-925-065A-23102	Sequence 23102, A
129	13.8	81.2	411	4	US-09-925-065A-562677	Sequence 562677, A	c 202	13.8	81.2	735	5	US-09-925-065A-23102	Sequence 23102, A
130	13.8	81.2	411	4	US-09-925-065A-562678	Sequence 562678, A	c 203	13.8	81.2	735	5	US-09-925-065A-23102	Sequence 23102, A
131	13.8	81.2	411	5	US-09-925-065A-562676	Sequence 562676, A	c 204	13.8	81.2	735	5	US-09-925-065A-23102	Sequence 23102, A
132	13.8	81.2	411	5	US-09-925-065A-562677	Sequence 562677, A	c 205	13.8	81.2	735	12	US-10-301-480-124338	Sequence 124338, A
133	13.8	81.2	411	5	US-09-925-065A-562678	Sequence 562678, A	c 206	13.8	81.2	735	12	US-10-301-480-124339	Sequence 124339, A
134	13.8	81.2	417	13	US-11-097-143-32339	Sequence 32339, A	c 207	13.8	81.2	735	12	US-10-301-480-737747	Sequence 737747, A
135	13.8	81.2	421	8	US-10-767-701-27209	Sequence 27209, A	c 208	13.8	81.2	735	12	US-10-301-480-737748	Sequence 737748, A
136	13.8	81.2	422	9	US-10-425-115-96719	Sequence 96719, A	c 209	13.8	81.2	748	8	US-10-437-963-25789	Sequence 25789, A
137	13.8	81.2	431	9	US-10-425-115-98389	Sequence 88389, A	c 210	13.8	81.2	751	8	US-10-038-854-55	Sequence 55, Appl
138	13.8	81.2	435	7	US-10-156-761-4409	Sequence 4409, Ap	c 211	13.8	81.2	779	6	US-10-027-632-126863	Sequence 126863, A
139	13.8	81.2	437	8	US-10-767-701-18962	Sequence 18962, A	c 212	13.8	81.2	779	6	US-10-027-632-126864	Sequence 126864, A
140	13.8	81.2	459	10	US-10-467-857-913	Sequence 913, App	c 213	13.8	81.2	779	7	US-10-027-632-126863	Sequence 126863, A
141	13.8	81.2	526	3	US-09-918-995-3480	Sequence 3480, Ap	c 214	13.8	81.2	779	7	US-10-027-632-126864	Sequence 126864, A
142	13.8	81.2	538	8	US-10-437-963-29206	Sequence 29206, A	c 215	13.8	81.2	816	7	US-10-156-761-5108	Sequence 5108, Ap
143	13.8	81.2	542	4	US-09-925-065A-342574	Sequence 342574, A	c 216	13.8	81.2	829	9	US-10-425-115-70599	Sequence 70599, A
144	13.8	81.2	542	5	US-09-925-065A-342574	Sequence 342574, A	c 217	13.8	81.2	838	6	US-10-027-632-134259	Sequence 134259, A
145	13.8	81.2	543	7	US-10-253-165-405	Sequence 405, App	c 218	13.8	81.2	838	7	US-10-027-632-134259	Sequence 134259, A
146	13.8	81.2	543	12	US-10-301-480-415320	Sequence 415320, A	c 219	13.8	81.2	867	8	US-10-282-122A-31724	Sequence 31724, A
147	13.8	81.2	543	12	US-10-301-480-1028729	Sequence 1028729, A	c 220	13.8	81.2	883	9	US-10-425-115-2830	Sequence 2830, Ap
148	13.8	81.2	546	4	US-09-925-065A-514429	Sequence 514429, A	c 221	13.8	81.2	896	6	US-10-027-632-172258	Sequence 172258, A
149	13.8	81.2	546	5	US-09-925-065A-514429	Sequence 514429, A	c 222	13.8	81.2	896	7	US-10-027-632-172258	Sequence 172258, A
150	13.8	81.2	546	7	US-10-259-165-61	Sequence 51, Appl	c 223	13.8	81.2	966	7	US-10-369-493-44361	Sequence 44361, A
151	13.8	81.2	548	4	US-09-925-065A-526604	Sequence 526604, A	c 224	13.8	81.2	971	6	US-10-027-632-120315	Sequence 120315, A
152	13.8	81.2	548	4	US-09-925-065A-526605	Sequence 526605, A	c 225	13.8	81.2	971	6	US-10-027-632-120316	Sequence 120316, A
153	13.8	81.2	548	4	US-09-925-065A-526606	Sequence 526606, A	c 226	13.8	81.2	971	6	US-10-027-632-120317	Sequence 120317, A
154	13.8	81.2	548	5	US-09-925-065A-526604	Sequence 526604, A	c 227	13.8	81.2	971	7	US-10-027-632-120316	Sequence 120316, A
155	13.8	81.2	548	5	US-09-925-065A-526605	Sequence 526605, A	c 228	13.8	81.2	971	7	US-10-027-632-120317	Sequence 120317, A
156	13.8	81.2	548	5	US-09-925-065A-526606	Sequence 526606, A	c 229	13.8	81.2	971	7	US-10-027-632-120317	Sequence 120317, A
157	13.8	81.2	574	3	US-09-918-995-28393	Sequence 28393, A	c 230	13.8	81.2	984	12	US-10-301-480-599267	Sequence 599267, A
158	13.8	81.2	588	6	US-10-027-632-284195	Sequence 284195, A	c 231	13.8	81.2	984	12	US-10-301-480-1212676	Sequence 1212676, A
159	13.8	81.2	588	7	US-10-027-632-284195	Sequence 284195, A	c 232	13.8	81.2	991	12	US-10-301-480-599119	Sequence 599119, A
160	13.8	81.2	592	4	US-09-925-065A-278677	Sequence 278677, A	c 233	13.8	81.2	991	12	US-10-301-480-1212528	Sequence 1212528, A
161	13.8	81.2	592	5	US-09-925-065A-278677	Sequence 278677, A	c 234	13.8	81.2	1050	7	US-10-369-493-31646	Sequence 31646, A
162	13.8	81.2	592	12	US-10-301-480-355681	Sequence 355681, A	c 235	13.8	81.2	1056	7	US-10-369-493-40780	Sequence 40780, A
163	13.8	81.2	592	12	US-10-301-480-969090	Sequence 969090, A	c 236	13.8	81.2	1059	8	US-10-437-963-94293	Sequence 94293, A

c 237	13.8	81.2	1101	8	US-10-282-122A-30651	Sequence 30651, A	310	13.8	81.2	3743	9	US-10-741-600-511	Sequence 511, App
238	13.8	81.2	1111	10	US-10-507-884-9	Sequence 9, Appli	c 311	13.8	81.2	3814	3	US-09-919-033-254	Sequence 254, App
239	13.8	81.2	1147	7	US-10-172-118-841	Sequence 841, App	312	13.8	81.2	3930	7	US-10-187-975-101	Sequence 101, App
240	13.8	81.2	1147	8	US-10-342-887-841	Sequence 841, App	313	13.8	81.2	3995	9	US-10-741-600-514	Sequence 514, App
241	13.8	81.2	1168	9	US-10-679-761-77	Sequence 77, Appl	314	13.8	81.2	4174	9	US-10-741-600-513	Sequence 513, App
242	13.8	81.2	1168	10	US-10-943-194-77	Sequence 77, Appl	315	13.8	81.2	4251	9	US-10-741-600-512	Sequence 512, App
243	13.8	81.2	1245	4	US-09-925-065A-36167	Sequence 36167, A	316	13.8	81.2	4257	7	US-10-161-403-30	Sequence 30, Appl
244	13.8	81.2	1245	4	US-09-925-065A-36168	Sequence 36168, A	317	13.8	81.2	4257	7	US-10-161-408-22	Sequence 22, Appl
245	13.8	81.2	1245	5	US-09-925-065A-36167	Sequence 36167, A	318	13.8	81.2	4257	10	US-10-161-408-22	Sequence 22, Appl
246	13.8	81.2	1245	5	US-09-925-065A-36168	Sequence 36168, A	319	13.8	81.2	4257	13	US-10-161-408-22	Sequence 22, Appl
247	13.8	81.2	1245	12	US-10-301-480-137405	Sequence 137405, A	320	13.8	81.2	4257	16	US-11-006-076-30	Sequence 30, Appl
248	13.8	81.2	1245	12	US-10-301-480-137406	Sequence 137406, A	321	13.8	81.2	4346	7	US-10-161-403-113	Sequence 113, App
249	13.8	81.2	1245	12	US-10-301-480-137406	Sequence 137406, A	322	13.8	81.2	4346	7	US-10-161-408-26	Sequence 26, Appl
250	13.8	81.2	1245	12	US-10-301-480-137406	Sequence 137406, A	323	13.8	81.2	4346	10	US-10-161-408-26	Sequence 26, Appl
251	13.8	81.2	1267	7	US-10-310-154-225	Sequence 225, App	324	13.8	81.2	4346	13	US-11-006-076-113	Sequence 113, App
252	13.8	81.2	1267	10	US-10-732-923-239	Sequence 239, App	325	13.8	81.2	4346	16	US-11-082-154A-113	Sequence 113, App
c 253	13.8	81.2	1290	8	US-10-282-122A-32173	Sequence 32173, A	326	13.8	81.2	4381	8	US-10-471-450-37	Sequence 37, Appl
c 254	13.8	81.2	1294	3	US-09-918-889-3	Sequence 3, Appli	c 327	13.8	81.2	4679	7	US-10-291-172-551	Sequence 551, App
c 255	13.8	81.2	1294	3	US-09-918-889-3	Sequence 3, Appli	c 328	13.8	81.2	4679	8	US-10-221-278-551	Sequence 551, App
c 256	13.8	81.2	1322	9	US-10-425-115-70598	Sequence 70598, A	329	13.8	81.2	4725	8	US-10-437-963-76873	Sequence 76873, A
c 257	13.8	81.2	1384	9	US-10-739-930-2320	Sequence 2320, Ap	330	13.8	81.2	4852	13	US-11-097-143-12688	Sequence 12688, A
c 258	13.8	81.2	1442	3	US-09-918-889-2	Sequence 2, Appli	331	13.8	81.2	5043	9	US-10-790-455-5	Sequence 5, Appli
c 259	13.8	81.2	1442	3	US-09-918-889-2	Sequence 2, Appli	332	13.8	81.2	5043	9	US-10-811-1368-5	Sequence 5, Appli
c 260	13.8	81.2	1488	8	US-10-282-122A-11356	Sequence 11356, A	333	13.8	81.2	5043	9	US-10-940-315-5	Sequence 5, Appli
c 261	13.8	81.2	1502	8	US-10-424-599-111856	Sequence 111856, A	334	13.8	81.2	5043	10	US-10-950-050-5	Sequence 5, Appli
c 262	13.8	81.2	1599	10	US-10-737-082-165	Sequence 165, App	335	13.8	81.2	5043	13	US-11-068-155-5	Sequence 5, Appli
c 263	13.8	81.2	1599	10	US-10-765-790-165	Sequence 165, App	336	13.8	81.2	5043	15	US-11-193-750-5	Sequence 5, Appli
c 264	13.8	81.2	1616	8	US-10-731-988-32	Sequence 32, Appl	337	13.8	81.2	5049	10	US-10-745-237-123	Sequence 123, App
c 265	13.8	81.2	1635	9	US-10-781-142-4	Sequence 4, Appli	338	13.8	81.2	5049	13	US-11-097-143-26117	Sequence 26117, A
c 266	13.8	81.2	1635	9	US-10-781-142-4	Sequence 4, Appli	339	13.8	81.2	5101	9	US-10-790-455-4	Sequence 4, Appli
c 267	13.8	81.2	1635	11	US-10-978-927-11	Sequence 11, Appl	340	13.8	81.2	5101	9	US-10-811-1368-4	Sequence 4, Appli
c 268	13.8	81.2	1635	11	US-10-978-927-11	Sequence 11, Appl	341	13.8	81.2	5101	9	US-10-940-315-4	Sequence 4, Appli
c 269	13.8	81.2	1637	8	US-10-437-963-40944	Sequence 40944, A	342	13.8	81.2	5101	10	US-10-950-050-4	Sequence 4, Appli
c 270	13.8	81.2	1671	10	US-10-821-234-557	Sequence 657, App	343	13.8	81.2	5101	13	US-11-068-155-4	Sequence 4, Appli
c 271	13.8	81.2	1731	7	US-10-369-493-47151	Sequence 47151, A	344	13.8	81.2	5101	15	US-11-193-750-4	Sequence 4, Appli
c 272	13.8	81.2	1807	6	US-10-198-846-12889	Sequence 12889, A	345	13.8	81.2	5147	7	US-10-715-270-1	Sequence 1, Appli
c 273	13.8	81.2	1837	8	US-10-622-896-1	Sequence 1, Appli	346	13.8	81.2	5192	7	US-10-161-403-88	Sequence 88, Appl
c 274	13.8	81.2	1839	9	US-10-844-874-7	Sequence 7, Appli	347	13.8	81.2	5192	13	US-11-006-076-88	Sequence 88, Appl
c 275	13.8	81.2	1857	9	US-10-781-142-16	Sequence 16, Appl	348	13.8	81.2	5192	16	US-11-082-154A-88	Sequence 88, Appl
c 276	13.8	81.2	1857	10	US-10-887-553A-283	Sequence 283, App	c 349	13.8	81.2	5354	16	US-11-136-527-1889	Sequence 1889, Ap
c 277	13.8	81.2	1857	10	US-10-947-249-196	Sequence 196, App	350	13.8	81.2	5382	7	US-10-331-329-21	Sequence 21, Appl
c 278	13.8	81.2	1857	11	US-10-978-927-16	Sequence 16, Appl	351	13.8	81.2	5712	9	US-10-715-270-3	Sequence 3, Appli
c 279	13.8	81.2	1958	3	US-09-918-889-3	Sequence 1, Appli	c 352	13.8	81.2	5979	6	US-10-098-979-1	Sequence 1, Appli
c 280	13.8	81.2	1958	3	US-09-918-889-3	Sequence 1, Appli	c 353	13.8	81.2	6066	13	US-11-097-143-18787	Sequence 18787, A
c 281	13.8	81.2	1976	8	US-10-717-597-47	Sequence 47, Appl	c 354	13.8	81.2	6124	9	US-10-714-000-15	Sequence 15, Appl
c 282	13.8	81.2	1976	9	US-10-394-808-4	Sequence 4, Appli	355	13.8	81.2	6124	9	US-10-715-270-6	Sequence 6, Appli
c 283	13.8	81.2	1976	9	US-10-803-482-4	Sequence 4, Appli	356	13.8	81.2	6233	9	US-10-790-455-10	Sequence 10, Appl
c 284	13.8	81.2	1976	10	US-10-887-553A-602	Sequence 602, App	357	13.8	81.2	6233	9	US-10-811-1368-10	Sequence 10, Appl
c 285	13.8	81.2	2085	3	US-09-729-674-93	Sequence 93, Appl	358	13.8	81.2	6233	10	US-10-940-315-10	Sequence 10, Appl
c 286	13.8	81.2	2085	9	US-10-913-553-93	Sequence 93, Appl	359	13.8	81.2	6233	10	US-10-950-050-10	Sequence 10, Appl
c 287	13.8	81.2	2115	16	US-11-096-568A-18423	Sequence 18423, A	360	13.8	81.2	6233	13	US-11-068-155-10	Sequence 10, Appl
c 288	13.8	81.2	2120	7	US-10-085-959-208	Sequence 208, App	361	13.8	81.2	6233	15	US-11-193-750-10	Sequence 10, Appl
c 289	13.8	81.2	2129	9	US-10-357-930-30281	Sequence 30281, A	362	13.8	81.2	6279	7	US-10-182-329-1	Sequence 1, Appli
c 290	13.8	81.2	2165	10	US-10-956-157-2275	Sequence 2275, Ap	363	13.8	81.2	6280	8	US-10-672-764A-64	Sequence 64, Appl
c 291	13.8	81.2	2172	8	US-10-389-647-345	Sequence 345, App	364	13.8	81.2	6289	8	US-10-807-308-7	Sequence 7, Appli
c 292	13.8	81.2	2174	6	US-10-142-516-1	Sequence 1, Appli	365	13.8	81.2	6502	8	US-10-672-764A-41	Sequence 41, Appl
c 293	13.8	81.2	2207	16	US-11-288-720-7	Sequence 7, Appli	c 366	13.8	81.2	6515	16	US-11-074-373-45	Sequence 45, Appl
c 294	13.8	81.2	2219	7	US-10-369-493-34386	Sequence 34386, A	c 367	13.8	81.2	6730	13	US-11-075-234-1	Sequence 1, Appli
c 295	13.8	81.2	2226	8	US-10-437-963-88149	Sequence 88149, A	368	13.8	81.2	6782	9	US-10-714-000-17	Sequence 17, Appl
c 296	13.8	81.2	2232	9	US-10-425-115-38382	Sequence 38382, A	369	13.8	81.2	6782	9	US-10-715-270-5	Sequence 5, Appli
c 297	13.8	81.2	2417	13	US-11-097-143-32338	Sequence 32338, A	370	13.8	81.2	7033	8	US-10-678-816-5	Sequence 5, Appli
c 298	13.8	81.2	2541	7	US-10-369-493-34272	Sequence 34272, A	371	13.8	81.2	7122	16	US-11-074-373-46	Sequence 46, Appl
c 299	13.8	81.2	2732	8	US-10-425-114-29468	Sequence 29468, A	c 372	13.8	81.2	7180	8	US-10-381-153-14	Sequence 14, Appl
c 300	13.8	81.2	2755	8	US-10-424-598-118015	Sequence 118015, A	c 373	13.8	81.2	7201	8	US-10-600-790-20	Sequence 20, Appl
c 301	13.8	81.2	2787	13	US-11-027-143-12689	Sequence 12689, A	374	13.8	81.2	7201	10	US-10-872-859-54	Sequence 54, Appl
c 302	13.8	81.2	3083	6	US-10-027-632-116026	Sequence 116026, A	c 375	13.8	81.2	7231	10	US-10-872-598A-42	Sequence 42, Appl
c 303	13.8	81.2	3083	6	US-10-027-632-116026	Sequence 116027, A	c 376	13.8	81.2	7299	10	US-10-872-598A-43	Sequence 43, Appl
c 304	13.8	81.2	3083	7	US-10-027-632-116026	Sequence 116026, A	c 377	13.8	81.2	7429	9	US-10-873-332-61	Sequence 61, Appl
c 305	13.8	81.2	3083	7	US-10-027-632-116027	Sequence 116027, A	c 378	13.8	81.2	7438	7	US-10-359-050-11	Sequence 11, Appl
c 306	13.8	81.2	3360	13	US-11-097-143-18788	Sequence 18788, A	379	13.8	81.2	7438	7	US-10-014-099F-5	Sequence 5, Appli
c 307	13.8	81.2	3486	8	US-10-282-122A-33541	Sequence 33541, A	c 380	13.8	81.2	7456	10	US-10-872-598A-41	Sequence 41, Appl
c 308	13.8	81.2	3500	15	US-11-085-775-1	Sequence 1, Appli	381	13.8	81.2	7523	7	US-10-014-099F-79	Sequence 79, Appl
c 309	13.8	81.2	3636	7	US-10-094-749-957	Sequence 957, App	382	13.8	81.2	7608	7	US-10-014-099F-78	Sequence 78, Appl

383	13.8	81.2	7803	7	US-10-014-099F-82	Sequence 82, Appl	C 456	13.8	81.2	18476	10	US-10-765-790-109	Sequence 109, App
C 384	13.8	81.2	7840	3	US-09-957-458B-5	Sequence 5, Appl	457	13.8	81.2	21293	11	US-10-330-773-272	Sequence 272, App
385	13.8	81.2	8157	7	US-10-014-099F-83	Sequence 83, Appl	458	13.8	81.2	24678	9	US-10-741-600-17702	Sequence 17702, A
C 386	13.8	81.2	8454	13	US-11-097-143-26116	Sequence 26116, A	C 459	13.8	81.2	25760	6	US-09-999-121-13	Sequence 13, Appl
387	13.8	81.2	8540	2	US-08-431-644A-12	Sequence 12, Appl	460	13.8	81.2	28690	6	US-10-010-802-1	Sequence 1, Appl
C 388	13.8	81.2	8742	7	US-10-105-616-6	Sequence 6, Appl	C 461	13.8	81.2	32846	7	US-10-017-1601-1657	Sequence 1657, Ap
C 389	13.8	81.2	8852	3	US-09-957-458B-6	Sequence 6, Appl	C 462	13.8	81.2	32846	7	US-10-292-798-1321	Sequence 1321, Ap
390	13.8	81.2	9027	7	US-10-117-960-11	Sequence 11, Appl	C 463	13.8	81.2	40558	9	US-10-629-318-37	Sequence 37, Appl
391	13.8	81.2	9249	7	US-10-389-120-2	Sequence 2, Appl	C 464	13.8	81.2	46410	16	US-11-124-367A-5085	Sequence 5085, Ap
392	13.8	81.2	9320	9	US-10-471-065-20	Sequence 20, Appl	C 465	13.8	81.2	47196	10	US-10-741-600-17636	Sequence 17636, A
393	13.8	81.2	9482	10	US-10-888-961-4	Sequence 4, Appl	466	13.8	81.2	47196	10	US-10-995-561-13249	Sequence 13249, A
C 394	13.8	81.2	9669	3	US-09-764-869-2072	Sequence 2072, Ap	467	13.8	81.2	60815	6	US-10-087-192-52	Sequence 52, Appl
C 395	13.8	81.2	9669	6	US-10-091-504-2072	Sequence 2072, Ap	468	13.8	81.2	60815	6	US-09-987-722-205	Sequence 205, Appl
C 396	13.8	81.2	9669	7	US-10-227-577-2072	Sequence 2072, Ap	469	13.8	81.2	70419	7	US-10-004-113-40	Sequence 40, Appl
397	13.8	81.2	9737	7	US-10-331-329-22	Sequence 22, Appl	470	13.8	81.2	70779	6	US-10-087-192-1012	Sequence 1012, Ap
398	13.8	81.2	9737	7	US-10-331-329-23	Sequence 23, Appl	471	13.8	81.2	80420	8	US-10-322-281-42	Sequence 42, Appl
399	13.8	81.2	9737	7	US-10-331-329-28	Sequence 28, Appl	472	13.8	81.2	100000	16	US-11-124-367A-5024	Sequence 5024, Ap
400	13.8	81.2	9871	7	US-10-331-329-24	Sequence 24, Appl	C 473	13.8	81.2	100048	8	US-10-672-787-39	Sequence 39, Appl
401	13.8	81.2	10060	7	US-10-331-329-25	Sequence 25, Appl	C 474	13.8	81.2	109147	11	US-10-330-773-695	Sequence 695, App
402	13.8	81.2	10233	7	US-10-050-898-283	Sequence 283, App	C 475	13.8	81.2	116585	6	US-10-087-192-133	Sequence 133, App
403	13.8	81.2	10285	7	US-10-050-902-283	Sequence 283, App	C 476	13.8	81.2	130427	6	US-10-175-523-87	Sequence 87, Appl
404	13.8	81.2	10325	6	US-10-112-267-53	Sequence 53, Appl	C 477	13.8	81.2	130427	13	US-11-099-266-87	Sequence 87, Appl
405	13.8	81.2	10330	6	US-10-112-267-53	Sequence 53, Appl	C 478	13.8	81.2	154746	3	US-09-827-688-8	Sequence 8, Appl
406	13.8	81.2	10477	8	US-10-656-269-24	Sequence 24, Appl	479	13.8	81.2	168749	7	US-10-085-117-250	Sequence 250, App
407	13.8	81.2	10516	8	US-10-656-269-22	Sequence 22, Appl	480	13.8	81.2	168749	15	US-11-121-086-61	Sequence 61, Appl
408	13.8	81.2	10551	7	US-10-656-269-20	Sequence 20, Appl	481	13.8	81.2	173602	15	US-11-121-086-25	Sequence 25, Appl
409	13.8	81.2	10561	8	US-10-656-269-18	Sequence 18, Appl	482	13.8	81.2	173602	15	US-11-121-086-104	Sequence 104, App
410	13.8	81.2	10615	8	US-10-656-269-45	Sequence 45, Appl	483	13.8	81.2	221000	7	US-10-174-014-12	Sequence 12, Appl
C 411	13.8	81.2	10774	13	US-11-075-234-3	Sequence 3, Appl	484	13.8	81.2	233380	6	US-10-087-192-652	Sequence 652, App
412	13.8	81.2	10774	8	US-10-656-269-23	Sequence 23, Appl	C 485	13.8	81.2	257645	11	US-10-330-773-266	Sequence 266, App
413	13.8	81.2	10921	8	US-10-656-269-21	Sequence 21, Appl	C 486	13.8	81.2	326002	10	US-10-461-862-56	Sequence 56, Appl
414	13.8	81.2	10961	8	US-10-656-269-19	Sequence 19, Appl	C 487	13.8	81.2	1694969	12	US-10-506-454-1690	Sequence 1690, Ap
415	13.8	81.2	11006	8	US-10-656-269-17	Sequence 17, Appl	488	13.8	81.2	9025608	7	US-10-156-761-1	Sequence 1, Appl
416	13.8	81.2	11059	8	US-10-656-269-46	Sequence 46, Appl	C 489	13.8	81.2	9025608	7	US-10-156-761-1	Sequence 1, Appl
417	13.8	81.2	11105	8	US-10-163-863A-32	Sequence 32, Appl	C 490	13.6	80.0	553	12	US-10-301-480-336638	Sequence 336638,
C 418	13.8	81.2	11110	9	US-10-781-142-20	Sequence 20, Appl	C 491	13.6	80.0	555	12	US-10-301-480-950047	Sequence 950047,
C 419	13.8	81.2	11110	11	US-10-978-827-20	Sequence 20, Appl	C 492	13.6	80.0	555	5	US-09-925-065A-258010	Sequence 258010,
420	13.8	81.2	11160	8	US-10-163-863A-8	Sequence 8, Appl	C 493	13.6	80.0	555	5	US-09-925-065A-258010	Sequence 258010,
421	13.8	81.2	11646	7	US-10-117-960-10	Sequence 10, Appl	C 494	13.4	78.8	19	11	US-10-310-914A-158192	Sequence 158192,
422	13.8	81.2	11941	8	US-10-163-863A-30	Sequence 30, Appl	C 495	13.4	78.8	20	11	US-10-310-914A-158221	Sequence 158221,
423	13.8	81.2	11998	8	US-10-163-863A-28	Sequence 28, Appl	C 496	13.4	78.8	20	11	US-10-310-914A-567181	Sequence 567181,
424	13.8	81.2	12041	7	US-10-117-960-9	Sequence 9, Appl	C 497	13.4	78.8	72	11	US-10-310-914A-5736	Sequence 5736, Ap
425	13.8	81.2	12088	8	US-10-163-863A-4	Sequence 4, Appl	C 498	13.4	78.8	72	11	US-10-310-914A-16410	Sequence 16410, A
C 426	13.8	81.2	12263	9	US-10-781-142-19	Sequence 19, Appl	C 499	13.4	78.8	72	11	US-10-310-914A-19214	Sequence 19214, A
C 427	13.8	81.2	12263	11	US-10-978-827-19	Sequence 19, Appl	500	13.4	78.8	90	11	US-10-310-914A-271	Sequence 271, App
428	13.8	81.2	12514	9	US-10-714-000-16	Sequence 16, Appl	C 501	13.4	78.8	130	7	US-10-029-386-21689	Sequence 21689, A
429	13.8	81.2	12514	9	US-10-715-270-4	Sequence 4, Appl	C 502	13.4	78.8	143	8	US-10-437-963-31423	Sequence 31423, A
430	13.8	81.2	12588	8	US-10-163-863A-27	Sequence 27, Appl	C 503	13.4	78.8	179	9	US-10-425-115-17045	Sequence 17045, A
431	13.8	81.2	12591	8	US-10-163-863A-7	Sequence 7, Appl	C 504	13.4	78.8	182	8	US-10-424-599-135303	Sequence 135303,
432	13.8	81.2	12680	8	US-10-163-863A-3	Sequence 3, Appl	C 505	13.4	78.8	201	10	US-10-995-561-64148	Sequence 64148, A
433	13.8	81.2	12753	8	US-10-041-018-19	Sequence 19, Appl	C 506	13.4	78.8	201	10	US-10-995-561-80338	Sequence 80338, A
434	13.8	81.2	13547	7	US-10-117-960-2	Sequence 2, Appl	C 507	13.4	78.8	201	10	US-10-995-561-80351	Sequence 80351, A
C 435	13.8	81.2	13766	8	US-10-105-616-1	Sequence 1, Appl	C 508	13.4	78.8	210	3	US-09-983-965-5890	Sequence 5890, Ap
C 436	13.8	81.2	14262	7	US-10-163-863A-9	Sequence 9, Appl	C 509	13.4	78.8	228	8	US-10-242-535A-46881	Sequence 46881, A
C 437	13.8	81.2	14987	10	US-10-931-080B-2	Sequence 2, Appl	C 510	13.4	78.8	228	8	US-10-085-783A-46881	Sequence 46881, A
438	13.8	81.2	15083	10	US-10-986-498-27	Sequence 27, Appl	C 511	13.4	78.8	274	8	US-10-085-783A-3819	Sequence 3819, Ap
439	13.8	81.2	15110	10	US-10-986-498-27	Sequence 27, Appl	C 512	13.4	78.8	274	8	US-10-085-783A-3819	Sequence 3819, Ap
C 440	13.8	81.2	17290	3	US-09-999-121-7	Sequence 7, Appl	C 513	13.4	78.8	275	9	US-10-425-115-40769	Sequence 40769, Ap
441	13.8	81.2	17344	13	US-11-097-143-27418	Sequence 27418, A	C 514	13.4	78.8	285	3	US-09-864-408A-8967	Sequence 8967, Ap
442	13.8	81.2	17358	13	US-11-097-143-12700	Sequence 12700, A	C 515	13.4	78.8	315	9	US-10-425-115-13828	Sequence 13828, A
443	13.8	81.2	17402	9	US-10-790-455-8	Sequence 8, Appl	C 516	13.4	78.8	317	9	US-10-425-115-160057	Sequence 160057,
444	13.8	81.2	17402	9	US-10-811-136B-8	Sequence 8, Appl	C 517	13.4	78.8	320	9	US-10-425-115-69870	Sequence 69870, A
445	13.8	81.2	17402	9	US-10-940-315-8	Sequence 8, Appl	C 518	13.4	78.8	327	8	US-10-437-963-80693	Sequence 80693, A
446	13.8	81.2	17402	10	US-10-950-050-8	Sequence 8, Appl	C 519	13.4	78.8	348	3	US-09-814-353-511	Sequence 511, App
447	13.8	81.2	17402	13	US-11-068-155-8	Sequence 8, Appl	C 520	13.4	78.8	348	3	US-09-814-353-6889	Sequence 6889, Ap
448	13.8	81.2	17402	15	US-11-193-750-8	Sequence 8, Appl	C 521	13.4	78.8	351	8	US-10-437-963-58539	Sequence 58539, A
449	13.8	81.2	18116	9	US-10-790-455-7	Sequence 7, Appl	C 522	13.4	78.8	369	9	US-10-437-963-32081	Sequence 32081, A
450	13.8	81.2	18116	9	US-10-811-136B-7	Sequence 7, Appl	C 523	13.4	78.8	394	3	US-10-425-115-85795	Sequence 85795, A
451	13.8	81.2	18116	9	US-10-940-315-7	Sequence 7, Appl	C 524	13.4	78.8	402	3	US-09-854-133-706	Sequence 706, App
452	13.8	81.2	18116	10	US-10-950-050-7	Sequence 7, Appl	C 525	13.4	78.8	402	6	US-10-144-649A-706	Sequence 706, App
C 453	13.8	81.2	18116	13	US-11-068-155-7	Sequence 7, Appl	C 526	13.4	78.8	407	8	US-10-424-599-25830	Sequence 25830, A
454	13.8	81.2	18116	15	US-11-193-750-7	Sequence 7, Appl	C 527	13.4	78.8	414	3	US-09-814-353-13274	Sequence 13274, A
C 455	13.8	81.2	18476	10	US-10-737-082-109	Sequence 109, App	C 528	13.4	78.8	451	6	US-10-027-632-126492	Sequence 126492,

529	13.4	78.8	7 US-10-027-632-126492	Sequence 126492,	c 602	13.4	78.8	722	6 US-10-027-632-150490	Sequence 150490,
530	13.4	78.8	8 US-10-027-632-126492	Sequence 65, Appl	c 603	13.4	78.8	722	6 US-10-027-632-150491	Sequence 150491,
531	13.4	78.8	8 US-10-027-632-126492	Sequence 4776, Ap	c 604	13.4	78.8	722	6 US-10-027-632-150489	Sequence 150489,
532	13.4	78.8	3 US-09-867-550-1533	Sequence 1533, Ap	c 605	13.4	78.8	722	7 US-10-027-632-150490	Sequence 150490,
533	13.4	78.8	4 US-09-925-065A-476158	Sequence 476158,	c 606	13.4	78.8	722	7 US-10-027-632-150491	Sequence 150491,
534	13.4	78.8	4 US-09-925-065A-476159	Sequence 476159,	c 607	13.4	78.8	742	4 US-09-925-065A-704419	Sequence 704419,
535	13.4	78.8	4 US-09-925-065A-476160	Sequence 476160,	c 608	13.4	78.8	742	5 US-09-925-065A-704419	Sequence 704419,
536	13.4	78.8	5 US-09-925-065A-476158	Sequence 476158,	c 609	13.4	78.8	759	7 US-10-264-049-1287	Sequence 1287, Ap
537	13.4	78.8	5 US-09-925-065A-476159	Sequence 476159,	c 610	13.4	78.8	771	16 US-11-096-568A-30720	Sequence 30720, A
538	13.4	78.8	5 US-09-925-065A-476160	Sequence 476160,	c 611	13.4	78.8	777	6 US-10-027-632-135035	Sequence 135035,
539	13.4	78.8	3 US-09-814-353-15584	Sequence 15584, A	c 612	13.4	78.8	777	6 US-10-027-632-135036	Sequence 135036,
540	13.4	78.8	8 US-10-021-323-4459	Sequence 4459, Ap	c 613	13.4	78.8	777	6 US-10-027-632-135037	Sequence 135037,
541	13.4	78.8	8 US-10-029-386-5341	Sequence 5341, Ap	c 614	13.4	78.8	777	6 US-10-027-632-135035	Sequence 135035,
542	13.4	78.8	7 US-10-029-386-7989	Sequence 7989, Ap	c 615	13.4	78.8	777	7 US-10-027-632-135036	Sequence 135036,
543	13.4	78.8	4 US-09-925-065A-367555	Sequence 367555,	c 616	13.4	78.8	777	7 US-10-027-632-135037	Sequence 135037,
544	13.4	78.8	5 US-09-925-065A-367555	Sequence 367555,	c 617	13.4	78.8	783	8 US-10-437-963-4779	Sequence 4779, Ap
545	13.4	78.8	6 US-10-027-632-284738	Sequence 284738,	c 618	13.4	78.8	804	8 US-10-437-963-4780	Sequence 4780, Ap
546	13.4	78.8	7 US-10-027-632-284738	Sequence 284738,	c 619	13.4	78.8	806	16 US-11-096-568A-4145	Sequence 4145, Ap
547	13.4	78.8	9 US-10-425-115-157628	Sequence 157628,	c 620	13.4	78.8	807	4 US-09-925-065A-8647	Sequence 8647, Ap
548	13.4	78.8	4 US-09-925-065A-740861	Sequence 740861,	c 621	13.4	78.8	807	4 US-09-925-065A-8648	Sequence 8648, Ap
549	13.4	78.8	5 US-09-925-065A-740861	Sequence 740861,	c 622	13.4	78.8	807	4 US-09-925-065A-8649	Sequence 8649, Ap
550	13.4	78.8	5 US-10-425-115-168017	Sequence 168017,	c 623	13.4	78.8	807	5 US-09-925-065A-8647	Sequence 8647, Ap
551	13.4	78.8	9 US-10-425-115-121720	Sequence 121720,	c 624	13.4	78.8	807	5 US-09-925-065A-8648	Sequence 8648, Ap
552	13.4	78.8	12 US-10-301-480-437868	Sequence 437868,	c 625	13.4	78.8	807	5 US-09-925-065A-8649	Sequence 8649, Ap
553	13.4	78.8	12 US-10-301-480-437868	Sequence 1051277,	c 626	13.4	78.8	807	12 US-10-301-480-109884	Sequence 109884,
554	13.4	78.8	4 US-09-925-065A-790875	Sequence 790875,	c 627	13.4	78.8	807	12 US-10-301-480-109885	Sequence 109885,
555	13.4	78.8	5 US-09-925-065A-790875	Sequence 790875,	c 628	13.4	78.8	807	12 US-10-301-480-109886	Sequence 109886,
556	13.4	78.8	4 US-09-925-065A-278676	Sequence 278676,	c 629	13.4	78.8	807	12 US-10-301-480-723293	Sequence 723293,
557	13.4	78.8	5 US-09-925-065A-278676	Sequence 278676,	c 630	13.4	78.8	807	12 US-10-301-480-723294	Sequence 723294,
558	13.4	78.8	12 US-10-301-480-355680	Sequence 355680,	c 631	13.4	78.8	807	12 US-10-301-480-723295	Sequence 723295,
559	13.4	78.8	12 US-10-301-480-969089	Sequence 969089,	c 632	13.4	78.8	818	6 US-10-027-632-163893	Sequence 163893,
560	13.4	78.8	3 US-09-974-300-5866	Sequence 5866, Ap	c 633	13.4	78.8	818	7 US-10-027-632-163893	Sequence 163893,
561	13.4	78.8	601 4 US-09-969-034-957	Sequence 957, App	c 634	13.4	78.8	835	7 US-10-029-386-24083	Sequence 24083, A
562	13.4	78.8	4 US-09-925-065A-213895	Sequence 213895,	c 635	13.4	78.8	846	16 US-11-096-568A-28049	Sequence 28049, A
563	13.4	78.8	5 US-09-925-065A-213895	Sequence 213895,	c 636	13.4	78.8	858	3 US-09-892-316-2	Sequence 2, Appl
564	13.4	78.8	12 US-10-301-480-300769	Sequence 300769,	c 637	13.4	78.8	902	8 US-10-437-963-69392	Sequence 69392, A
565	13.4	78.8	12 US-10-301-480-914178	Sequence 914178,	c 638	13.4	78.8	915	8 US-10-425-114-13026	Sequence 13026, A
566	13.4	78.8	6 US-10-027-632-260608	Sequence 260608,	c 639	13.4	78.8	921	6 US-10-027-632-33411	Sequence 33411, A
567	13.4	78.8	7 US-10-027-632-260608	Sequence 260608,	c 640	13.4	78.8	921	7 US-10-027-632-33411	Sequence 33411, A
568	13.4	78.8	12 US-10-301-480-95420	Sequence 95420, A	c 641	13.4	78.8	922	8 US-10-767-701-7773	Sequence 7773, Ap
569	13.4	78.8	12 US-10-301-480-708829	Sequence 708829,	c 642	13.4	78.8	926	8 US-10-767-701-10411	Sequence 10411, A
570	13.4	78.8	4 US-09-925-065A-915893	Sequence 915893,	c 643	13.4	78.8	936	8 US-10-425-114-16138	Sequence 16138, A
571	13.4	78.8	5 US-09-925-065A-915893	Sequence 915893,	c 644	13.4	78.8	936	10 US-10-644-765-72	Sequence 72, Appl
572	13.4	78.8	3 US-09-969-034-914	Sequence 914, App	c 645	13.4	78.8	974	8 US-10-425-114-21847	Sequence 21847, A
573	13.4	78.8	4 US-09-925-065A-915101	Sequence 915101,	c 646	13.4	78.8	984	8 US-10-425-114-21847	Sequence 21847, A
574	13.4	78.8	5 US-09-925-065A-915101	Sequence 915101,	c 647	13.4	78.8	986	16 US-11-096-568A-19160	Sequence 19160, A
575	13.4	78.8	12 US-10-301-480-46970	Sequence 46970, A	c 648	13.4	78.8	988	12 US-10-301-480-600510	Sequence 600510,
576	13.4	78.8	12 US-10-301-480-660379	Sequence 660379,	c 649	13.4	78.8	988	12 US-10-301-480-1213919	Sequence 1213919,
577	13.4	78.8	4 US-09-925-065A-864559	Sequence 864559,	c 650	13.4	78.8	1029	4 US-09-925-065A-697154	Sequence 697154,
578	13.4	78.8	5 US-09-925-065A-864559	Sequence 864559,	c 651	13.4	78.8	1029	4 US-09-925-065A-697155	Sequence 697155,
579	13.4	78.8	16 US-11-120-665-4	Sequence 4, Appl	c 652	13.4	78.8	1029	4 US-09-925-065A-697156	Sequence 697156,
580	13.4	78.8	4 US-09-925-065A-514467	Sequence 514467,	c 653	13.4	78.8	1029	4 US-09-925-065A-697157	Sequence 697157,
581	13.4	78.8	5 US-09-925-065A-514467	Sequence 514467,	c 654	13.4	78.8	1029	5 US-09-925-065A-697154	Sequence 697154,
582	13.4	78.8	4 US-09-925-065A-863199	Sequence 863199,	c 655	13.4	78.8	1029	5 US-09-925-065A-697155	Sequence 697155,
583	13.4	78.8	6 US-09-925-065A-863199	Sequence 863199,	c 656	13.4	78.8	1029	5 US-09-925-065A-697156	Sequence 697156,
584	13.4	78.8	9 US-10-425-115-137152	Sequence 137152,	c 657	13.4	78.8	1029	5 US-09-925-065A-697157	Sequence 697157,
585	13.4	78.8	12 US-10-301-480-536884	Sequence 536884,	c 658	13.4	78.8	1041	9 US-10-425-115-19734	Sequence 19734, A
586	13.4	78.8	12 US-10-301-480-1150293	Sequence 1150293,	c 659	13.4	78.8	1075	8 US-10-437-963-66970	Sequence 66970, A
587	13.4	78.8	6 US-10-027-632-189161	Sequence 189161,	c 660	13.4	78.8	1118	4 US-09-925-065A-31673	Sequence 31673, A
588	13.4	78.8	7 US-10-027-632-189162	Sequence 189162,	c 661	13.4	78.8	1118	4 US-09-925-065A-31674	Sequence 31674, A
589	13.4	78.8	6 US-10-027-632-189161	Sequence 189161,	c 662	13.4	78.8	1118	5 US-09-925-065A-31673	Sequence 31673, A
590	13.4	78.8	7 US-10-027-632-189162	Sequence 189162,	c 663	13.4	78.8	1118	5 US-09-925-065A-31674	Sequence 31674, A
591	13.4	78.8	4 US-09-925-065A-939940	Sequence 939940,	c 664	13.4	78.8	1118	12 US-10-301-480-132911	Sequence 132911,
592	13.4	78.8	5 US-09-925-065A-939940	Sequence 939940,	c 665	13.4	78.8	1118	12 US-10-301-480-132912	Sequence 132912,
593	13.4	78.8	6 US-10-027-632-113470	Sequence 113470,	c 666	13.4	78.8	1118	12 US-10-301-480-746320	Sequence 746320,
594	13.4	78.8	6 US-10-027-632-113471	Sequence 113471,	c 667	13.4	78.8	1118	12 US-10-301-480-746321	Sequence 746321,
595	13.4	78.8	6 US-10-027-632-113472	Sequence 113472,	c 668	13.4	78.8	1176	9 US-10-282-122A-6858	Sequence 6858, Ap
596	13.4	78.8	6 US-10-027-632-113470	Sequence 113470,	c 669	13.4	78.8	1176	9 US-10-841-316-1	Sequence 1, Appl
597	13.4	78.8	7 US-10-027-632-113471	Sequence 113471,	c 670	13.4	78.8	1176	9 US-10-841-316-2	Sequence 2, Appl
598	13.4	78.8	7 US-10-027-632-113472	Sequence 113472,	c 671	13.4	78.8	1203	9 US-10-858-706-1	Sequence 1, Appl
599	13.4	78.8	9 US-10-425-115-91933	Sequence 91933, A	c 672	13.4	78.8	1203	9 US-10-858-706-3	Sequence 3, Appl
600	13.4	78.8	719 9 US-10-425-115-19736	Sequence 19736, A	c 673	13.4	78.8	1217	7 US-10-341-434-160	Sequence 160, App
601	13.4	78.8	6 US-10-027-632-150489	Sequence 150489,	c 674	13.4	78.8	1217	7 US-10-341-434-164	Sequence 164, App

675	13.4	78.8	1217	10	US-10-489-740-56	Sequence 56, Appl	748	13.4	78.8	1803	5	US-09-925-065A-35171	Sequence 35171, A
c 676	13.4	78.8	1224	8	US-10-282-122A-33401	Sequence 33401, A	749	13.4	78.8	1803	12	US-10-301-480-136409	Sequence 136409, A
c 677	13.4	78.8	1229	8	US-10-425-114-35303	Sequence 33303, A	750	13.4	78.8	1803	12	US-10-301-480-749818	Sequence 749818, A
c 678	13.4	78.8	1234	9	US-10-425-115-103197	Sequence 103197, A	751	13.4	78.8	1814	7	US-10-184-771-5	Sequence 5, Appl
c 679	13.4	78.8	1239	9	US-10-425-115-77929	Sequence 77929, A	752	13.4	78.8	1814	7	US-10-186-047-5	Sequence 5, Appl
c 680	13.4	78.8	1260	7	US-10-156-761-3588	Sequence 3588, A	753	13.4	78.8	1814	8	US-10-644-187-5	Sequence 5, Appl
c 681	13.4	78.8	1395	8	US-10-282-122A-19398	Sequence 1398, A	754	13.4	78.8	1814	9	US-10-926-720-5	Sequence 5, Appl
c 682	13.4	78.8	1400	10	US-10-956-157-6027	Sequence 6027, Ap	755	13.4	78.8	1814	10	US-10-980-759-5	Sequence 5, Appl
c 683	13.4	78.8	1400	13	US-11-060-756-2261	Sequence 2261, Ap	756	13.4	78.8	1814	13	US-11-064-196-5	Sequence 5, Appl
c 684	13.4	78.8	1400	13	US-11-060-756-6533	Sequence 6533, Ap	757	13.4	78.8	1815	9	US-10-425-115-26787	Sequence 26787, A
c 685	13.4	78.8	1419	7	US-10-081-872-87	Sequence 87, Appl	758	13.4	78.8	1815	10	US-10-425-115-26787	Sequence 1, Appl
c 686	13.4	78.8	1419	7	US-10-105-733-3	Sequence 3, Appl	759	13.4	78.8	1815	9	US-10-473-092-1	Sequence 12, Appl
c 687	13.4	78.8	1419	8	US-10-081-739A-3	Sequence 3, Appl	760	13.4	78.8	1815	9	US-10-433-579-12	Sequence 23836, A
c 688	13.4	78.8	1419	8	US-10-385-305-87	Sequence 87, Appl	761	13.4	78.8	1815	9	US-10-357-930-23836	Sequence 29742, A
c 689	13.4	78.8	1419	10	US-10-489-510-3	Sequence 3, Appl	762	13.4	78.8	1815	13	US-11-007-819-9	Sequence 9, Appl
c 690	13.4	78.8	1419	16	US-11-096-568A-25129	Sequence 25129, A	763	13.4	78.8	1815	13	US-11-007-819-13	Sequence 13, Appl
c 691	13.4	78.8	1425	7	US-10-369-493-28649	Sequence 28649, A	764	13.4	78.8	1815	9	US-10-425-115-154128	Sequence 154128, A
c 692	13.4	78.8	1449	7	US-10-369-493-31407	Sequence 31407, A	765	13.4	78.8	1815	9	US-09-925-065A-64270	Sequence 64270, A
c 693	13.4	78.8	1482	3	US-09-822-830A-46	Sequence 46, Appl	766	13.4	78.8	1815	5	US-09-925-065A-64270	Sequence 64270, A
c 694	13.4	78.8	1524	10	US-10-750-185-58349	Sequence 58349, A	767	13.4	78.8	1815	12	US-10-301-480-165508	Sequence 165508, A
c 695	13.4	78.8	1524	10	US-10-750-623-58349	Sequence 58349, A	768	13.4	78.8	1815	12	US-10-301-480-178917	Sequence 778917, A
c 696	13.4	78.8	1527	8	US-10-437-963-95923	Sequence 95923, A	769	13.4	78.8	1815	7	US-10-104-047-56	Sequence 56, Appl
c 697	13.4	78.8	1548	3	US-09-769-864-11	Sequence 11, Appl	770	13.4	78.8	1815	16	US-11-072-512-56	Sequence 56, Appl
c 698	13.4	78.8	1548	3	US-09-854-346-5	Sequence 5, Appl	771	13.4	78.8	1815	10	US-10-750-185-53793	Sequence 53793, A
c 699	13.4	78.8	1548	3	US-09-902-188A-6	Sequence 6, Appl	772	13.4	78.8	1815	10	US-10-750-623-53793	Sequence 53793, A
c 700	13.4	78.8	1548	3	US-09-918-543-5	Sequence 5, Appl	773	13.4	78.8	1815	7	US-10-108-260A-559	Sequence 559, App
c 701	13.4	78.8	1548	3	US-09-925-576C-5	Sequence 5, Appl	774	13.4	78.8	1815	8	US-10-425-114-34374	Sequence 34374, A
c 702	13.4	78.8	1548	6	US-10-146-327-7	Sequence 7, Appl	775	13.4	78.8	1815	8	US-10-425-115-65985	Sequence 65985, A
c 703	13.4	78.8	1548	6	US-10-025-648-6	Sequence 6, Appl	776	13.4	78.8	1815	10	US-10-956-157-2468	Sequence 2468, Ap
c 704	13.4	78.8	1548	7	US-10-327-837-11	Sequence 11, Appl	777	13.4	78.8	1815	8	US-10-741-601-261	Sequence 261, Ap
c 705	13.4	78.8	1548	8	US-10-665-667-11	Sequence 11, Appl	778	13.4	78.8	1815	9	US-10-741-600-745	Sequence 745, App
c 706	13.4	78.8	1548	8	US-10-477-725-5	Sequence 5, Appl	779	13.4	78.8	1815	9	US-10-425-115-19738	Sequence 19738, A
c 707	13.4	78.8	1548	9	US-10-779-418-6	Sequence 6, Appl	780	13.4	78.8	1815	10	US-10-926-188A-34	Sequence 34, Appl
c 708	13.4	78.8	1548	10	US-10-778-469-6	Sequence 6, Appl	781	13.4	78.8	1815	3	US-09-880-107-3690	Sequence 3690, Ap
c 709	13.4	78.8	1548	10	US-10-980-923-11	Sequence 11, Appl	782	13.4	78.8	1815	3	US-09-968-007A-436	Sequence 436, App
c 710	13.4	78.8	1548	10	US-10-630-203-5	Sequence 5, Appl	783	13.4	78.8	1815	7	US-10-099-322-39	Sequence 39, Appl
c 711	13.4	78.8	1548	11	US-10-779-418-6	Sequence 6, Appl	784	13.4	78.8	1815	8	US-10-044-564-39	Sequence 39, Appl
c 712	13.4	78.8	1548	13	US-11-039-565-11	Sequence 11, Appl	785	13.4	78.8	1815	10	US-10-843-641A-6906	Sequence 6906, Ap
c 713	13.4	78.8	1548	15	US-11-103-037-6	Sequence 6, Appl	786	13.4	78.8	1815	11	US-10-826-585-25	Sequence 25, Appl
c 714	13.4	78.8	1549	7	US-10-190-377-8	Sequence 8, Appl	787	13.4	78.8	1815	16	US-11-054-281-39	Sequence 39, Appl
c 715	13.4	78.8	1569	9	US-10-425-115-41062	Sequence 41062, A	788	13.4	78.8	1815	9	US-10-006-285-484	Sequence 484, App
c 716	13.4	78.8	1615	3	US-09-820-002-1	Sequence 1, Appl	789	13.4	78.8	1815	7	US-10-739-930-4146	Sequence 4146, Ap
c 717	13.4	78.8	1615	7	US-10-274-031-1	Sequence 1, Appl	790	13.4	78.8	1815	13	US-11-097-143-35341	Sequence 35341, A
c 718	13.4	78.8	1615	15	US-11-182-752-1	Sequence 1, Appl	791	13.4	78.8	1815	10	US-10-450-763-7019	Sequence 7019, Ap
c 719	13.4	78.8	1626	16	US-11-102-188-8	Sequence 8, Appl	792	13.4	78.8	1815	10	US-10-858-730-39	Sequence 39, Appl
c 720	13.4	78.8	1641	3	US-09-738-626-2676	Sequence 2676, Ap	793	13.4	78.8	1815	13	US-11-097-143-12970	Sequence 12970, A
c 721	13.4	78.8	1650	7	US-10-081-872-91	Sequence 91, Appl	794	13.4	78.8	1815	13	US-11-097-143-35320	Sequence 35320, A
c 722	13.4	78.8	1650	7	US-10-081-872-103	Sequence 103, App	795	13.4	78.8	1815	8	US-10-466-085A-6	Sequence 6, Appl
c 723	13.4	78.8	1650	8	US-10-385-305-91	Sequence 91, Appl	796	13.4	78.8	1815	16	US-11-120-665-1	Sequence 1, Appl
c 724	13.4	78.8	1650	8	US-10-385-305-103	Sequence 103, App	797	13.4	78.8	1815	9	US-10-370-715B-65	Sequence 65, Appl
c 725	13.4	78.8	1650	16	US-11-102-188-1	Sequence 1, Appl	798	13.4	78.8	1815	10	US-10-750-185-27792	Sequence 27792, A
c 726	13.4	78.8	1664	4	US-09-925-065A-86578	Sequence 86578, A	799	13.4	78.8	1815	10	US-10-750-623-27792	Sequence 27792, A
c 727	13.4	78.8	1664	5	US-09-925-065A-86578	Sequence 86578, A	800	13.4	78.8	1815	13	US-11-097-143-6548	Sequence 6548, Ap
c 728	13.4	78.8	1664	12	US-10-301-480-187818	Sequence 187818, A	801	13.4	78.8	1815	13	US-10-437-963-26530	Sequence 26530, A
c 729	13.4	78.8	1664	12	US-10-301-480-801227	Sequence 801227, A	802	13.4	78.8	1815	7	US-10-108-260A-950	Sequence 950, App
c 730	13.4	78.8	1763	13	US-11-082-389-181	Sequence 181, App	803	13.4	78.8	1815	13	US-11-097-143-34918	Sequence 34918, A
c 731	13.4	78.8	1783	3	US-09-919-048-188	Sequence 188, App	804	13.4	78.8	1815	6	US-10-027-632-114266	Sequence 114266, A
c 732	13.4	78.8	1783	3	US-09-776-191-67	Sequence 67, Appl	805	13.4	78.8	1815	6	US-10-027-632-114267	Sequence 114267, A
c 733	13.4	78.8	1783	6	US-10-102-283-188	Sequence 188, App	806	13.4	78.8	1815	7	US-10-027-632-114266	Sequence 114266, A
c 734	13.4	78.8	1783	6	US-10-073-060-1	Sequence 1, Appl	807	13.4	78.8	1815	7	US-10-027-632-114267	Sequence 114267, A
c 735	13.4	78.8	1783	6	US-10-135-795-188	Sequence 188, App	808	13.4	78.8	1815	3	US-09-927-827-22	Sequence 22, Appl
c 736	13.4	78.8	1783	6	US-10-205-823-177	Sequence 177, App	809	13.4	78.8	1815	7	US-10-406-209-48	Sequence 18, Appl
c 737	13.4	78.8	1783	7	US-10-210-120-1	Sequence 1, Appl	810	13.4	78.8	1815	13	US-11-097-143-926	Sequence 926, App
c 738	13.4	78.8	1783	7	US-10-295-027-1160	Sequence 1160, Ap	811	13.4	78.8	1815	9	US-10-425-115-108158	Sequence 108158, A
c 739	13.4	78.8	1783	7	US-10-156-214A-34	Sequence 34, Appl	812	13.4	78.8	1815	13	US-11-097-143-6547	Sequence 6547, Ap
c 740	13.4	78.8	1783	8	US-10-652-932-188	Sequence 188, App	813	13.4	78.8	1815	8	US-10-466-085A-32	Sequence 32, Appl
c 741	13.4	78.8	1783	10	US-10-956-157-792	Sequence 792, App	814	13.4	78.8	1815	9	US-10-855-897-2	Sequence 2, Appl
c 742	13.4	78.8	1783	10	US-10-909-035-1	Sequence 1, Appl	815	13.4	78.8	1815	3	US-09-924-101-2	Sequence 2, Appl
c 743	13.4	78.8	1783	13	US-11-051-454-177	Sequence 177, App	816	13.4	78.8	1815	6	US-09-764-847-1077	Sequence 1077, Ap
c 744	13.4	78.8	1783	15	US-11-112-908-14	Sequence 14, Appl	817	13.4	78.8	1815	6	US-10-092-154-1077	Sequence 1077, Ap
c 745	13.4	78.8	1783	15	US-11-000-688-380	Sequence 380, App	818	13.4	78.8	1815	3	US-09-845-064-17	Sequence 17, Appl
c 746	13.4	78.8	1798	8	US-10-425-114-22306	Sequence 22306, A	819	13.4	78.8	1815	3	US-09-845-064-19	Sequence 19, Appl
c 747	13.4	78.8	1803	4	US-09-925-065A-35171	Sequence 35171, A	820	13.4	78.8	1815	16	US-11-173-792-9	Sequence 9, Appl

Sequence 525399,
Sequence 525400,
Sequence 1138805,
Sequence 1138806,
Sequence 1138807,
Sequence 1138808,
Sequence 1138809,
Sequence 388662,
Sequence 1002071,
Sequence 4465, Ap
Sequence 700567,
Sequence 700567,
Sequence 509347,
Sequence 509347,
Sequence 97, Appl
Sequence 555185,
Sequence 1168594,
Sequence 98, Appl
Sequence 98, Appl
Sequence 80, Appl
Sequence 58342, A
Sequence 95, Appl
Sequence 95, Appl
Sequence 173732,
Sequence 173732,
Sequence 96, Appl
Sequence 96, Appl
Sequence 38867, A
Sequence 39485, A
Sequence 39117, A
Sequence 81, Appl
Sequence 81, Appl

ALIGNMENTS

RESULT 1
US-10-600-816-31
; Sequence 31, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-31

Query Match 100.0%; Score 17; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGAG 17
| | | | | | | | | | | | | | | | | |
Db 1 GCCCACGCTTGCCGAG 17

RESULT 2

US-10-600-816-30
; Sequence 30, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-30

Query Match 100.0%; Score 17; DB 8; Length 42;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGAG 17
| | | | | | | | | | | | | | | | | |
Db 1 GCCCACGCTTGCCGAG 17

RESULT 3
US-10-066-543-1937
; Sequence 1937, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1937
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1937

Query Match 100.0%; Score 17; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGAG 17
| | | | | | | | | | | | | | | | | |
Db 344 GCCCACGCTTGCCGAG 360

RESULT 4
US-10-066-543-1811/c
; Sequence 1811, Application US/10066543

Publication No. US20030087818A1

GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Darick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
PRIORITY FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1811
LENGTH: 552
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1811

Query Match 100.0%; Score 17; DB 6; Length 552;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||

DB 154 GCCCAGCTTGGCCGAG 138

RESULT 5

US-09-969-034-2222/c
Sequence 2222, Application US/09969034
Publication No. US20040110668A1
GENERAL INFORMATION:
APPLICANT: Burgess, Christopher C.
APPLICANT: Astle, Jon H.
APPLICANT: Carroll, Eddie III
APPLICANT: Catino, Theodore J.
APPLICANT: Dwivedi, Poornima
APPLICANT: Molino, Gary A.
APPLICANT: Thiagalingam, Arunthathi
APPLICANT: Lewis, Marcia E.
TITLE OF INVENTION: Nucleic Acid Sequences Differentially
FILE REFERENCE: 1657/1032
CURRENT APPLICATION NUMBER: US/09/969,034
CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/237,271
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 4494
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2222
LENGTH: 620
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 491, 500, 517, 530, 545, 547, 553, 558, 570, 599, 620
OTHER INFORMATION: n = A,T,C or G
US-09-969-034-2222

Query Match 100.0%; Score 17; DB 3; Length 620;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||

DB 154 GCCCAGCTTGGCCGAG 138

RESULT 6

US-09-969-034-3393/c
Sequence 3393, Application US/09969034
Publication No. US20040110668A1
GENERAL INFORMATION:
APPLICANT: Burgess, Christopher C.
APPLICANT: Astle, Jon H.
APPLICANT: Carroll, Eddie III
APPLICANT: Catino, Theodore J.
APPLICANT: Dwivedi, Poornima
APPLICANT: Molino, Gary A.
APPLICANT: Thiagalingam, Arunthathi
APPLICANT: Lewis, Marcia E.
TITLE OF INVENTION: Nucleic Acid Sequences Differentially
FILE REFERENCE: 1657/1032
CURRENT APPLICATION NUMBER: US/09/969,034
CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/237,271
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 4494
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3393
LENGTH: 634
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 363, 470, 472, 516, 523, 531, 551, 567, 573, 588, 594, 613,
LOCATION: 623
OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3393

Query Match 100.0%; Score 17; DB 3; Length 634;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||

DB 158 GCCCAGCTTGGCCGAG 142

RESULT 7

US-10-125-968-701/c
Sequence 701, Application US/10125968
Publication No. US20030215805A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Paterno, Adam
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
APPLICANT: Elias, Josh
APPLICANT: Mertens, Maureen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-032
CURRENT APPLICATION NUMBER: US/10/125,968
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/285,163
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 1417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 701
LENGTH: 642
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 72, 76, 77, 99, 101, 423, 445, 468, 489, 557, 566, 633

Query Match 100.0%; Score 17; DB 3; Length 620;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||

```
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-701

Query Match      100.0%; Score 17; DB 7; Length 642;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 184 GCCCAGCGTTGCCGAG 168

RESULT 8
US-10-712-615-134
; Sequence 134, Application US/10712615
; Publication No. US20040214317A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV8, EXPRESSED
; FILE REFERENCE: D0047A-CIP
; CURRENT APPLICATION NUMBER: US/10/712,615
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: U.S. 09/992,238
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: U.S. 60/248,285
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: U.S. 60/268,581
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: U.S. 60/308,285
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: U.S. 60/317,166
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 134
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: dna,Homo sapiens
US-10-712-615-134

Query Match      100.0%; Score 17; DB 9; Length 1071;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 1018 GCCCAGCGTTGCCGAG 1034

RESULT 9
US-09-866-050A-249
; Sequence 249, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-866-050A-249
```

```
Query Match      100.0%; Score 17; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 990 GCCCAGCGTTGCCGAG 1006

RESULT 10
US-10-152-661-249
; Sequence 249, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-10-152-661-249

Query Match      100.0%; Score 17; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGCCGAG 17
Db 990 GCCCAGCGTTGCCGAG 1006

RESULT 11
US-10-313-542-223
; Sequence 223, Application US/10313542
; Publication No. US20030120057A1
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/10/313,542
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/495,050
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: 1999-02-01
```

; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030120057A1 2681738CTI
US-10-313-542-223

Query Match 100.0%; Score 17; DB 7; Length 1228;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
|||
Db 541 GCCCAGCTTGCCGAG 557

RESULT 12

US-11-060-756-2418
; Sequence 2418, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2418
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-2418

Query Match 100.0%; Score 17; DB 13; Length 1400;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
|||
Db 215 GCCCAGCTTGCCGAG 231

RESULT 13

US-11-060-756-6690
; Sequence 6690, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6690
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-6690

Query Match 100.0%; Score 17; DB 13; Length 1400;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
|||
Db 215 GCCCAGCTTGCCGAG 231

RESULT 14

US-10-936-626-64
; Sequence 64, Application US/10936626
; Publication No. US20050106644A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-64

Query Match 100.0%; Score 17; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
|||
Db 1145 GCCCAGCTTGCCGAG 1161

RESULT 15

US-10-938-061-64
; Sequence 64, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul

APPLICANT: Spencer, Susan D.
APPLICANT: Smith, Victoria
APPLICANT: Williams, P. Mickey
APPLICANT: Wu, Thomas D.
APPLICANT: Zhang, Zemin
APPLICANT: Sakanaka, Chie
APPLICANT: Chuntharapai, Anan
APPLICANT: Reed Chae J.
TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
TREATMENT OF Tumor
FILE REFERENCE: P5001RLP1B
CURRENT APPLICATION NUMBER: US/10/938,061
CURRENT FILING DATE: 2004-09-10
PRIOR APPLICATION NUMBER: US 10/872,991
PRIOR FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: US 10/872,972
PRIOR FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: US 10/241,220
PRIOR FILING DATE: 2002-09-11
PRIOR APPLICATION NUMBER: US 10/177,488
PRIOR FILING DATE: 2002-06-19
PRIOR APPLICATION NUMBER: US 60/299,500
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 60/301,880
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/323,268
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US 60/557,116
PRIOR FILING DATE: 2004-03-26
PRIOR APPLICATION NUMBER: US 60/598,899
PRIOR FILING DATE: 2004-08-04
NUMBER OF SEQ ID NOS: 154
SEQ ID NO 64
LENGTH: 1460
TYPE: DNA
ORGANISM: Homo sapiens
US-10-938-061-64

Query Match 100.0%; Score 17; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1145 GCCCAGCGTTGGCCGAG 1161

RESULT 16
US-10-224-289-5
Sequence 5, Application US/10224289
Publication No. US20030207288A1
GENERAL INFORMATION:
APPLICANT: LEWIN, DAVID A.
APPLICANT: STEWART, TIMOTHY A.
TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
METHOD OF INVENTION: NUCLEIC ACID
FILE REFERENCE: 9800081-0085
CURRENT APPLICATION NUMBER: US/10/224,289
CURRENT FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/313,940
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapiens
US-10-224-289-5

Query Match 100.0%; Score 17; DB 7; Length 1619;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1137 GCCCAGCGTTGGCCGAG 1153

RESULT 17
US-10-935-190-43
Sequence 43, Application US/10935190
Publication No. US20050037466A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: AU-YOUNG, Janice
APPLICANT: BANDMAN, Olga
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: AZIMZAI, Valda
APPLICANT: BURFORD, Neil
APPLICANT: BAUGHN, Mariah R.
APPLICANT: LU, Dyoung Aina M.
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: PATTERSON, Chandra
APPLICANT: LAL, Preeti
TITLE OF INVENTION: RECEPTORS AND ASSOCIATED PROTEINS
FILE REFERENCE: PE-0726 PCT
CURRENT APPLICATION NUMBER: US/10/935,190
CURRENT FILING DATE: 2004-09-08
PRIOR APPLICATION NUMBER: US/10/031,904
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: 60/145,232; 60/158,578; 60/165,192
PRIOR FILING DATE: 1999-07-21; 1999-10-07; 1999-11-12
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PERL Program
SEQ ID NO 43
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 2681738CB1
US-10-935-190-43

Query Match 100.0%; Score 17; DB 9; Length 1619;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 1137 GCCCAGCGTTGGCCGAG 1153

RESULT 18
US-10-505-486-196
Sequence 196, Application US/10505486
Publication No. US20050118639A1
GENERAL INFORMATION:
APPLICANT: Takeda Chemical Industries, Ltd.
TITLE OF INVENTION: Determination of a ligand
FILE REFERENCE: P03-0006PCT
CURRENT APPLICATION NUMBER: US/10/505,486
CURRENT FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: JP 2002-45728
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: JP 2002-213949
PRIOR FILING DATE: 2002-07-23
PRIOR APPLICATION NUMBER: JP 2002-298237
PRIOR FILING DATE: 2002-10-11
NUMBER OF SEQ ID NOS: 233
SEQ ID NO 196
LENGTH: 1788
TYPE: DNA
ORGANISM: Human
US-10-505-486-196

```
Query Match      100.0%; Score 17; DB 10; Length 1788;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGCCCGAG 17
      |||
Db      1018 GCCCAGCTTGCCCGAG 1034

RESULT 19
US-10-775-920-13
; Sequence 13, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 00108
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-13

Query Match      100.0%; Score 17; DB 9; Length 2297;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGCCCGAG 17
      |||
Db      1125 GCCCAGCTTGCCCGAG 1141

RESULT 20
US-10-224-289-3
; Sequence 3, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; TITLE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-3

Query Match      100.0%; Score 17; DB 7; Length 2302;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGCCCGAG 17
      |||
Db      1117 GCCCAGCTTGCCCGAG 1133

RESULT 21
US-10-240-425-405
```

```
; Sequence 405, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scheff, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 405
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF095448
US-10-240-425-405

Query Match      100.0%; Score 17; DB 8; Length 2302;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGCCCGAG 17
      |||
Db      1117 GCCCAGCTTGCCCGAG 1133

RESULT 22
US-10-775-920-9
; Sequence 9, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 00108
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-9

Query Match      100.0%; Score 17; DB 9; Length 2302;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGCCCGAG 17
      |||
Db      1117 GCCCAGCTTGCCCGAG 1133

RESULT 23
US-10-510-507-2
; Sequence 2, Application US/10510507
; Publication No. US20050282165A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Terrett, Jonathan A
; TITLE OF INVENTION: DIAGNOSIS OF CARCINOMA USING RAIG1 POLYPEPTIDES
; FILE REFERENCE: 2543-1-039PCT/US
; CURRENT APPLICATION NUMBER: US/10/510,507
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: GB0208331.9
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: GB0221538.2
; PRIOR FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-510-507-2

Query Match      100.0%; Score 17; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCCGAG 17
      |||
Db      1117 GCCCAGCGCTTGCCCGAG 1133

RESULT 24
US-10-775-920-12
; Sequence 12, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: MerGen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: MerGen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-12

Query Match      100.0%; Score 17; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCCGAG 17
      |||
Db      1121 GCCCAGCGCTTGCCCGAG 1137

RESULT 25
US-10-176-847-59
; Sequence 59, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-59

Query Match      100.0%; Score 17; DB 6; Length 2316;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCCGAG 17
      |||
Db      1123 GCCCAGCGCTTGCCCGAG 1139

RESULT 26
US-11-080-991-59
; Sequence 59, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-080-991-59

Query Match      100.0%; Score 17; DB 15; Length 2316;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCCGAG 17
      |||
Db      1123 GCCCAGCGCTTGCCCGAG 1139

RESULT 27
US-10-775-920-11
; Sequence 11, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: MerGen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: MerGen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 2446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-11

Query Match      100.0%; Score 17; DB 9; Length 2446;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCCCAGCGCTTGCCCGAG 17
      |||
Db      1271 GCCCAGCGCTTGCCCGAG 1287
```

RESULT 28
US-10-225-567A-453
; Sequence 453, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 453
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-453

Query Match 100.0%; Score 17; DB 6; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
Db 1271 GCCCAGCTTGGCCGAG 1287

RESULT 29
US-10-269-909-63
; Sequence 63, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-63

Query Match 100.0%; Score 17; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
Db 1271 GCCCAGCTTGGCCGAG 1287

RESULT 30
US-10-269-909-64
; Sequence 64, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-64

Query Match 100.0%; Score 17; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-64

Query Match 100.0%; Score 17; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
Db 1271 GCCCAGCTTGGCCGAG 1287

RESULT 31
US-10-295-027-619
; Sequence 619, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500JUS
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 619

```
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-619

Query Match      100.0%; Score 17; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGCGAG 17
Db 1271 GCCCACGCTTGCCGCGAG 1287

RESULT 32
US-10-600-816-2
; Sequence 2, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RA13, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-2

Query Match      100.0%; Score 17; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGCGAG 17
Db 1271 GCCCACGCTTGCCGCGAG 1287

RESULT 33
US-10-600-816-18
; Sequence 18, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RA13, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RA13 Polymorphic Allele Summary Sequence.
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (112)..(112)
; OTHER INFORMATION: wherein "n" equals either G or A.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (364)..(364)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1111)..(1111)
; OTHER INFORMATION: wherein "n" equals either T or C.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1173)..(1173)
; OTHER INFORMATION: wherein "n" equals either A or G.
US-10-600-816-18

Query Match      100.0%; Score 17; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGCGAG 17
Db 1271 GCCCACGCTTGCCGCGAG 1287

RESULT 34
US-10-775-920-10
; Sequence 10, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-10

Query Match      100.0%; Score 17; DB 9; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCGCGAG 17
Db 1271 GCCCACGCTTGCCGCGAG 1287

RESULT 35
```

US-10-936-626-40
; Sequence 40, Application US/10936626
; Publication No. US20050106644A1
; GENERAL INFORMATION:

; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor

; FILE REFERENCE: P5001RIP1
; CURRENT APPLICATION NUMBER: US/10/936,626
; PRIOR FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 40

; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-40

Query Match 100.0%; Score 17; DB 10; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCCAGCTTGGCCGAG 17
Db 1271 GCCCAGCTTGGCCGAG 1287

RESULT 36

US-10-938-061-40
; Sequence 40, Application US/10938061
; Publication No. US20050107595A1

; GENERAL INFORMATION:

; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie

; APPLICANT: Chuntharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor

; FILE REFERENCE: P5001RIP1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; PRIOR FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 40
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-938-061-40

Query Match 100.0%; Score 17; DB 10; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCCAGCTTGGCCGAG 17
Db 1271 GCCCAGCTTGGCCGAG 1287

RESULT 37

US-11-169-041-32
; Sequence 32, Application US/11169041
; Publication No. US20060019284A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens

US-11-169-041-32

Query Match 100.0%; Score 17; DB 16; Length 2456;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCCAGCTTGGCCGAG 17
Db 1271 GCCCAGCTTGGCCGAG 1287

```
RESULT 38
US-10-264-049-834
; Sequence 834, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 834
; LENGTH: 2593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-049-834

Query Match 100.0%; Score 17; DB 7; Length 2593;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGCGAG 17
Db 1383 GCCCAGCTTGCCGCGAG 1399

RESULT 39
US-10-198-846-10424
; Sequence 10424, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10424
; LENGTH: 4239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 4238, 4239
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10424

Query Match 100.0%; Score 17; DB 6; Length 4239;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGCGAG 17
Db 1679 GCCCAGCTTGCCGCGAG 1695

RESULT 40
US-10-335-977-1640/c
; Sequence 1640, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 1640:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1020 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1020
; SEQUENCE DESCRIPTION: SEQ ID NO: 1640:
US-10-335-977-1640

Query Match 90.6%; Score 15.4; DB 8; Length 1020;
Best Local Similarity 94.1%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGCGAG 17
Db 981 GCCCAGCTTGCGCGAG 965

RESULT 41
US-10-335-977-1641/c
; Sequence 1641, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
```


ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1641:
SEQUENCE CHARACTERISTICS:
LENGTH: 1050 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1050
SEQUENCE DESCRIPTION: SEQ ID NO: 1641:

US-10-335-977-1641

Query Match 90.6%; Score 15.4; DB 8; Length 1050;
Best Local Similarity 94.1%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
DB 1014 GCCCAGCTTGGCCGAG 998

RESULT 42

US-10-956-157-41
; Sequence 41, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:

; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-41

Query Match 90.6%; Score 15.4; DB 10; Length 1173;
Best Local Similarity 94.1%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
DB 228 GCCCAGCTTGGCCGAG 244

RESULT 43

US-10-956-157-5276
; Sequence 5276, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:

; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5276
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-5276

Query Match 90.6%; Score 15.4; DB 10; Length 1173;
Best Local Similarity 94.1%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
DB 228 GCCCAGCTTGGCCGAG 244

RESULT 44

US-10-398-221-3404/c
; Sequence 3404, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:

; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3404
; LENGTH: 2306
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3404

Query Match 90.6%; Score 15.4; DB 8; Length 2306;
Best Local Similarity 94.1%; Pred. No. 2.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGAG 17
|||||
DB 834 GCCCAGCTTGGCCGAG 818

RESULT 45

US-09-757-716-2
; Sequence 2, Application US/09757716
; Patent No. US20010012515A1
; GENERAL INFORMATION:

; APPLICANT: Magna, Holly
; APPLICANT: Schaffer, Paul

```
; APPLICANT: Lawton, Michael
; APPLICANT: Yocum, Sue
; APPLICANT: Mitchell, Peter
; APPLICANT: Hutchinson, Nancy
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHOHYDROLASE-2
; FILE REFERENCE: PF-0420 US
; CURRENT APPLICATION NUMBER: US/09/757,716
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US/08/996,083
; PRIOR FILING DATE: 1997-12-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 4183
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4183)
; OTHER INFORMATION: n = A,T,C or G
; PUBLICATION INFORMATION:
US-09-757-716-2

Query Match          90.6%; Score 15.4; DB 3; Length 4183;
Best Local Similarity 94.1%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 3254 GCCCAGCGTTGGCCGAG 3270

RESULT 46
US-10-723-860-7430
; Sequence 7430, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziiz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7430
; LENGTH: 4222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2438)..(2503)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7430

Query Match          90.6%; Score 15.4; DB 9; Length 4222;
Best Local Similarity 94.1%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 3295 GCCCAGCGTTGGCCGAG 3311

RESULT 47
US-10-322-281-557
; Sequence 557, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 557
; LENGTH: 63761
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(63761)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-557

Query Match          90.6%; Score 15.4; DB 8; Length 63761;
Best Local Similarity 94.1%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 34522 GCCCAGCGTTGGCCGAG 34538

RESULT 48
US-10-322-281-560
; Sequence 560, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 560
; LENGTH: 68732
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(68732)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-560

Query Match          90.6%; Score 15.4; DB 8; Length 68732;
Best Local Similarity 94.1%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 38569 GCCCAGCGTTGGCCGAG 38585

RESULT 49
US-10-398-221-8
; Sequence 8, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
```

```
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 557
; LENGTH: 63761
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(63761)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-557

Query Match          90.6%; Score 15.4; DB 8; Length 63761;
Best Local Similarity 94.1%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 34522 GCCCAGCGTTGGCCGAG 34538

RESULT 48
US-10-322-281-560
; Sequence 560, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 560
; LENGTH: 68732
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(68732)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-560

Query Match          90.6%; Score 15.4; DB 8; Length 68732;
Best Local Similarity 94.1%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGAG 17
Db 38569 GCCCAGCGTTGGCCGAG 38585

RESULT 49
US-10-398-221-8
; Sequence 8, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
```

;; PRIOR FILING DATE: 2000-10-04
;; NUMBER OF SEQ ID NOS: 4025
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 8
;; LENGTH: 495269
;; TYPE: DNA
;; ORGANISM: Listeria innocua
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(end)
;; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-8

Query Match 90.6%; Score 15.4; DB 8; Length 495269;
Best Local Similarity 94.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
Db 400813 GCCCAGCTTGCCGAG 400829

RESULT 50

US-10-470-565-1/c
;; Sequence 1, Application US/10470565
;; Publication No. US20040126870A1
;; GENERAL INFORMATION:
;; APPLICANT: Societe des Produits Nestle S.A.
;; TITLE OF INVENTION: NC2705 - the genome of a Bifidobacterium
;; FILE REFERENCE: 80290/WO
;; CURRENT APPLICATION NUMBER: US/10/470,565
;; CURRENT FILING DATE: 2003-07-29
;; PRIOR APPLICATION NUMBER: EP 01102050.0
;; PRIOR FILING DATE: 2001-01-30
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 2256646
;; TYPE: DNA
;; ORGANISM: Bifidobacterium longum
US-10-470-565-1

Query Match 90.6%; Score 15.4; DB 8; Length 2256646;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
Db 1222897 GCCCAGCTTGCCGAG 1222881

RESULT 51

US-11-045-004-1/c
;; Sequence 1, Application US/11045004
;; Publication No. US20060078901A1
;; GENERAL INFORMATION:
;; APPLICANT: BUCHRIESER, CARMEN
;; APPLICANT: FRANGEUL, LIONEL
;; APPLICANT: COUVE, ELISABETH
;; APPLICANT: RUSNIOK, CHRISTOPHE
;; APPLICANT: FSIHI, HAFIDA
;; APPLICANT: DEHOUX, PIERRE
;; APPLICANT: DUSSURGET, OLIVIER
;; APPLICANT: CHETOUANI, FARID
;; APPLICANT: NEJARI, HAFED
;; APPLICANT: GLASER, PHILIPPE
;; APPLICANT: KUNST, FRANK
;; APPLICANT: COSSART, PASCALE
;; APPLICANT: DANIELS, JUSTIN
;; APPLICANT: GOEBEL, WERNER
;; APPLICANT: KREFT, JURGEN
;; APPLICANT: KUHN, MICHAEL
;; APPLICANT: NG, EVA

;; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
;; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
;; APPLICANT: GARRIDO-GARCIA, PATRICIA
;; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
;; APPLICANT: AMEND, ALEXANDRA
;; APPLICANT: CHAKRABORTY, TRINAD
;; APPLICANT: DOMANN, EUGEN
;; APPLICANT: HAIN, THORSTEN
;; APPLICANT: BERCHE, PATRICK
;; APPLICANT: CHARBIT, ALAIN
;; APPLICANT: DURANT, LIONEL
;; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
;; APPLICANT: BAQUERO, FERNANDO
;; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
;; APPLICANT: GOMEZ-LOPEZ, NURIA
;; APPLICANT: MADUENIO, ENCARNIA
;; APPLICANT: PABLOS, BETRIZ DE
;; APPLICANT: WEHLAND, JURGEN
;; APPLICANT: KARST, UWE
;; APPLICANT: ENTIAN, KARL-DIETER
;; APPLICANT: HAUF, JORG
;; APPLICANT: ROSE, MATTHIAS
;; APPLICANT: VOSS, HAMUT
;; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
;; FILE REFERENCE: 05394.0018-02
;; CURRENT APPLICATION NUMBER: US/11/045,004
;; CURRENT FILING DATE: 2005-01-28
;; PRIOR APPLICATION NUMBER: 10/637,657
;; PRIOR FILING DATE: 2003-08-11
;; PRIOR APPLICATION NUMBER: 10/257,023
;; PRIOR FILING DATE: 2002-10-08
;; PRIOR APPLICATION NUMBER: PCT/FR01/01118
;; PRIOR FILING DATE: 2001-04-11
;; PRIOR APPLICATION NUMBER: FR 00/04,629
;; PRIOR FILING DATE: 2000-04-11
;; NUMBER OF SEQ ID NOS: 2854
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 1
;; LENGTH: 2944528
;; TYPE: DNA
;; ORGANISM: Listeria monocytogenes
US-11-045-004-1

Query Match 90.6%; Score 15.4; DB 16; Length 2944528;
Best Local Similarity 94.1%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGAG 17
Db 1351366 GCCCAGCTTGCTGAG 1351350

RESULT 52

US-10-398-221-2058/c
;; Sequence 2058, Application US/10398221
;; Publication No. US20040018514A1
;; GENERAL INFORMATION:
;; APPLICANT: KUNST, Frederik
;; APPLICANT: GLASER, Philippe
;; TITLE OF INVENTION: Listeria innocua, genome and applications
;; FILE REFERENCE: 344 702 - US
;; CURRENT APPLICATION NUMBER: US/10/398,221
;; CURRENT FILING DATE: 2003-03-27
;; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
;; PRIOR FILING DATE: 2001-10-04
;; PRIOR APPLICATION NUMBER: FR 00/12 697
;; PRIOR FILING DATE: 2000-10-04
;; NUMBER OF SEQ ID NOS: 4025
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 2058
;; LENGTH: 3011208
;; TYPE: DNA
;; ORGANISM: Listeria innocua

US-10-398-221-2058

```
Query Match      90.6%; Score 15.4; DB 8; Length 3011208;
Best Local Similarity 94.1%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0;
```

Qy 1 GCCACGCTGGCCGAG 17
1358141 GCCACGCTGGCTGAG 1358125

RESULT 53

```

US-10-357-930-57611/C
; Sequence 57611, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57611
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-57611

```

```
Query Match      88.2%; Score 15; DB 9; Length 337;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 15: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 3 CCACGCTTGGCCGAG 17
|||
Dp 216 CCACGCTTGGCCGAG 202

RESULT 54

```

US-10-425-115-123425/C
; Sequence 123425, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21 (53222) B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 123425
; LENGTH: 404

```

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_44044C.1
US-10-425-115-123425

```

Query Match 88.2%; Score 15; DB 9; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels

Qy 1 GCCACGCTTGGCG 15
pb 136 GCCACGCTTGGCG 122

RESULT 55

```

US-10-739-930-4894/c
; Sequence 4894, Application US/10739930
; Publication No. US20040216190A1
;
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
;
; SEQ ID NO 4894
; LENGTH: 3303
; TYPE: DNA
; ORGANISM: Triticum aestivum
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER413830_1
US-10-739-930-4894

```

```
Query Match      88.2%; Score 15; DB 9; Length 3303;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 15: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 2 CCACGCTTGGCCGA 16
p6 619 CCACGCTTGGCCGA 605

RESULT 56

```

US-10-956-157-193814
; Sequence 193814, Application US/10956157
; Publication NO. US20050118625A1
; GENERAL INFORMATION:
;   APPLICANT: Wyeth
;   TITLE OF INVENTION: Mounts, William
;   TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION AND HUMAN PROTEASES
;   FILE REFERENCE: 031896-043000 (AM 101081)
;   CURRENT APPLICATION NUMBER: US/10/956.157
;   CURRENT FILING DATE: 2004-10-04
;   NUMBER OF SEQ ID NOS: 319805
;   SOFTWARE: Patentin version 3.2
;   SEQ ID NO 193814
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Probe Sequence
; US-10-956-157-193814

```

Query Match 84.7%; Score 14.4; DB 10; Length 25;
Best Local Similarity 93.8%; Pred. NO. 1.7e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy 2 CCCACGCTTGGCCGAG 17
|||
Db 1 CCCACGCTTGGCCCAAG 16
|||

```
RESULT 57
US-10-995-561-64090/c
; Sequence 64090, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64090
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-64090

Query Match      84.7%; Score 14.4; DB 10; Length 201;
Best Local Similarity 93.8%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGGCCGAG 17
      |||||
Db      158 CCCACGCTTGGCCGAG 143

RESULT 58
US-09-803-719-429/c
; Sequence 429, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 429
; LENGTH: 344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(344)
; OTHER INFORMATION: n = A,T,C or G
US-09-803-719-429
```

```
Query Match      84.7%; Score 14.4; DB 3; Length 344;
Best Local Similarity 93.8%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGGCCGAG 17
      |||||
Db      252 CCCACGCTTGGCCGAG 237

RESULT 59
US-10-779-543-13699/c
; Sequence 13699, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13699
; LENGTH: 344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 344
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-13699
```

```
Query Match      84.7%; Score 14.4; DB 10; Length 344;
Best Local Similarity 93.8%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGGCCGAG 17
      |||||
Db      252 CCCACGCTTGGCCGAG 237

RESULT 60
US-10-357-930-54041
; Sequence 54041, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
```

; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54041
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 28, 41, 42
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-54041

Query Match 84.7%; Score 14.4; DB 9; Length 570;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCAGCTTGCCGCA 16
Db 140 GCCCAGCTTGACCA 155
|||||

RESULT 61

US-09-925-065A-939940/c
; Sequence 939940, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 939940
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-939940

Query Match 84.7%; Score 14.4; DB 4; Length 653;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGAG 17

Db 648 CCCAGGCTTGCCGAG 633
|||||

RESULT 62

US-09-925-065A-939940/c
; Sequence 939940, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 939940
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-939940

Query Match 84.7%; Score 14.4; DB 5; Length 653;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGAG 17
Db 648 CCCAGGCTTGCCGAG 633
|||||

RESULT 63

US-09-925-065A-63251/c
; Sequence 63251, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63251
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-63251

Query Match 84.7%; Score 14.4; DB 4; Length 834;
Best Local Similarity 93.8%; Pred. No. 1e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
Db 69 CCCACCTTGCCGAG 54

RESULT 64
US-09-925-065A-63251/c
; Sequence 63251, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63251
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-63251

Query Match 84.7%; Score 14.4; DB 5; Length 834;
Best Local Similarity 93.8%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
Db 69 CCCACCTTGCCGAG 54

RESULT 65
US-10-301-480-164489/c
; Sequence 164489, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164489
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-164489

Query Match 84.7%; Score 14.4; DB 12; Length 834;
Best Local Similarity 93.8%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17

Db 69 CCCACCTTGCCGAG 54

RESULT 66
US-10-301-480-777898/c
; Sequence 777898, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 777898
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-777898

Query Match 84.7%; Score 14.4; DB 12; Length 834;
Best Local Similarity 93.8%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
Db 69 CCCACCTTGCCGAG 54

RESULT 67
US-10-101-510-469/c
; Sequence 469, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 469
; LENGTH: 1889
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-469

Query Match 84.7%; Score 14.4; DB 7; Length 1889;
Best Local Similarity 93.8%; Pred. No. 9.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCGAG 17
Db 18 CCCACGCTTGCCGAG 3

RESULT 68
US-09-861-846-3/c
; Sequence 3, Application US/09861846
; Patent No. US20020110852A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,


```

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001065
; CURRENT APPLICATION NUMBER: US/09/861,846
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/752,821
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13608
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(13608)
; OTHER INFORMATION: n = A,T,C or G
US-09-861-846-3

```

```

Query Match      84.7%; Score 14.4; DB 3; Length 13608;
Best Local Similarity 93.8%; Pred. No. 7.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      2 CCCAGCGTTGCCGAG 17
Db      7742 CCCAGCGTTGGCTGAG 7727

```

RESULT 69

```

US-10-250-463-3/c
; Sequence 3, Application US/10250463
; Publication No. US20040106775A1
; GENERAL INFORMATION:
; APPLICANT: PE CORPORATION (NY)
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001065
; CURRENT APPLICATION NUMBER: US/10/250,463
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: 09/752,821
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 09/861,846
; PRIOR FILING DATE: 2001-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(13608)
; OTHER INFORMATION: n = A,T,C or G
US-10-250-463-3

```

```

Query Match      84.7%; Score 14.4; DB 8; Length 13608;
Best Local Similarity 93.8%; Pred. No. 7.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      2 CCCAGCGTTGCCGAG 17
Db      7742 CCCAGCGTTGGCTGAG 7727

```

RESULT 70

```

US-10-294-934-485/c
; Sequence 485, Application US/10294934
; Publication No. US20040038231A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilva
; APPLICANT: Bougueleret, Lydie

```

```

; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62.US4.DIV
; CURRENT APPLICATION NUMBER: US/10/294,934
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/671,317
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm
; SEQ ID NO 485
; LENGTH: 49312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5466..7466
; OTHER INFORMATION: 5'regulatory region
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7467..7725
; OTHER INFORMATION: exon 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 20256..20355
; OTHER INFORMATION: exon 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 36905..36975
; OTHER INFORMATION: exon 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 45167..45248
; OTHER INFORMATION: exon 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 45728..45965
; OTHER INFORMATION: exon 5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 45966..49312
; OTHER INFORMATION: 3'regulatory region
; FEATURE:
; NAME/KEY: allele
; LOCATION: 7564
; OTHER INFORMATION: 10-286-289 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 7619
; OTHER INFORMATION: 10-286-345 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 7649
; OTHER INFORMATION: 10-286-375 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 17258
; OTHER INFORMATION: 12-425-57 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 21590
; OTHER INFORMATION: 12-421-135 : insertion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 21595
; OTHER INFORMATION: 12-421-140 : polymorphic base A or G

```

```
FEATURE:
NAME/KEY: allele
LOCATION: 36971
OTHER INFORMATION: 10-523-232 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 45214
OTHER INFORMATION: 10-289-201 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 45741
OTHER INFORMATION: 10-290-37 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 46029
OTHER INFORMATION: 10-290-326 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 46032
OTHER INFORMATION: 10-290-328 : deletion of G
FEATURE:
NAME/KEY: primer bind
LOCATION: 7276..7294
OTHER INFORMATION: 10-286.pu
FEATURE:
NAME/KEY: primer bind
LOCATION: 7676..7694
OTHER INFORMATION: 10-286.rp complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 16839..16856
OTHER INFORMATION: 12-425.rp
FEATURE:
NAME/KEY: primer bind
LOCATION: 17297..17314
OTHER INFORMATION: 12-425.pu complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 21456..21474
OTHER INFORMATION: 12-421.pu
FEATURE:
NAME/KEY: primer bind
LOCATION: 21886..21906
OTHER INFORMATION: 12-421.rp complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 36740..36758
OTHER INFORMATION: 10-523.pu
FEATURE:
NAME/KEY: primer bind
LOCATION: 36997..37015
OTHER INFORMATION: 10-523.rp complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 45020..45037
OTHER INFORMATION: 10-289.pu
FEATURE:
NAME/KEY: primer bind
LOCATION: 45413..45432
OTHER INFORMATION: 10-289.rp complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 45705..45724
OTHER INFORMATION: 10-290.pu
FEATURE:
NAME/KEY: primer bind
LOCATION: 46104..46123
OTHER INFORMATION: 10-290.rp complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 7545..7563
OTHER INFORMATION: 10-286-289.mis
FEATURE:
```

```
NAME/KEY: primer bind
LOCATION: 7565..7583
OTHER INFORMATION: 10-286-289.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 7600..7618
OTHER INFORMATION: 10-286-345.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 7620..7638
OTHER INFORMATION: 10-286-345.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 7630..7648
OTHER INFORMATION: 10-286-375.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 7650..7668
OTHER INFORMATION: 10-286-375.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 17239..17257
OTHER INFORMATION: 12-425-57.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 17259..17277
OTHER INFORMATION: 12-425-57.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 21576..21594
OTHER INFORMATION: 12-421-140.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 21596..21614
OTHER INFORMATION: 12-421-140.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 36952..36970
OTHER INFORMATION: 10-523-232.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 36972..36990
OTHER INFORMATION: 10-523-232.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 45195..45213
OTHER INFORMATION: 10-289-201.mis
FEATURE:
```

```
Query Match      84.7%  Score 14.4;  DB 8;  Length 49312;
Best Local Similarity 93.8%  Pred. No. 6.1e+02;
Matches 15;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;
```

```
Qy      2  CCCACGGTTGCCGAG 17
Db      12970  CCCACTCTGGCCGAG 12955
|||||
```

RESULT 71

```
US-11-204-311-485/c
; Sequence 485, Application US/11204311
; Publication No. US20060040304A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62 US4 DIV
; CURRENT APPLICATION NUMBER: US/11/204,311
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: US/10/294,934
; PRIOR FILING DATE: 2000-09-27
```

```

; PRIOR APPLICATION NUMBER: US 09/671,317
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm
; SEQ ID NO 485
; LENGTH: 49312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5466..7466
; OTHER INFORMATION: 5' regulatory region
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7467..7725
; OTHER INFORMATION: exon 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 20256..20355
; OTHER INFORMATION: exon 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 36905..36975
; OTHER INFORMATION: exon 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 45167..45248
; OTHER INFORMATION: exon 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 45728..45965
; OTHER INFORMATION: exon 5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 45966..49312
; OTHER INFORMATION: 3' regulatory region
; FEATURE:
; NAME/KEY: allele
; LOCATION: 7564
; OTHER INFORMATION: 10-286-289 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 7619
; OTHER INFORMATION: 10-286-345 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 7649
; OTHER INFORMATION: 10-286-375 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 17258
; OTHER INFORMATION: 12-425-57 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 21590
; OTHER INFORMATION: 12-421-135 : insertion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 21595
; OTHER INFORMATION: 12-421-140 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 36971
; OTHER INFORMATION: 10-523-232 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 45214
; OTHER INFORMATION: 10-289-201 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 45741
; OTHER INFORMATION: 10-290-37 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 46029
; OTHER INFORMATION: 10-290-326 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 46032
; OTHER INFORMATION: 10-290-328 : deletion of G
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 7276..7294
; OTHER INFORMATION: 10-286.pu
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 7676..7694
; OTHER INFORMATION: 10-286.rp complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 16839..16856
; OTHER INFORMATION: 12-425.rp
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 17297..17314
; OTHER INFORMATION: 12-425.pu complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 21456..21474
; OTHER INFORMATION: 12-421.pu
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 21886..21906
; OTHER INFORMATION: 12-421.rp complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 36740..36758
; OTHER INFORMATION: 10-523.pu
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 36997..37015
; OTHER INFORMATION: 10-523.rp complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 45020..45037
; OTHER INFORMATION: 10-289.pu
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 45705..45724
; OTHER INFORMATION: 10-290.pu
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 46104..46123
; OTHER INFORMATION: 10-290.rp complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 7545..7563
; OTHER INFORMATION: 10-286-289.mis
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 7565..7583
; OTHER INFORMATION: 10-286-289.mis complement
; FEATURE:
; NAME/KEY: primer_bind

```

LOCATION: 7600..7618
OTHER INFORMATION: 10-286-345.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 7630..7638
OTHER INFORMATION: 10-286-345.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 7630..7648
OTHER INFORMATION: 10-286-375.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 7650..7668
OTHER INFORMATION: 10-286-375.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 17239..17257
OTHER INFORMATION: 12-425-57.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 17259..17277
OTHER INFORMATION: 12-425-57.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 21576..21594
OTHER INFORMATION: 12-421-140.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 21596..21614
OTHER INFORMATION: 12-421-140.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 36952..36970
OTHER INFORMATION: 10-523-232.mis
FEATURE:
NAME/KEY: primer bind
LOCATION: 36972..36990
OTHER INFORMATION: 10-523-232.mis complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 45195..45213

Query Match 84.7%; Score 14.4; DB 16; Length 49312;
Best Local Similarity 93.8%; Pred. No. 6.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGAG 17
Db 12970 CCCACTCTTGCCGAG 12955

RESULT 72
US-10-461-862-56
; Sequence 56, Application US/10461862
; Publication No. US20050090434A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; CURRENT APPLICATION NUMBER: 529452001800
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 326002
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-461-862-56

Query Match 84.7%; Score 14.4; DB 10; Length 326002;
Best Local Similarity 93.8%; Pred. No. 4.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCGAG 17
Db 110753 CCCACCTTGCCGAG 110768
RESULT 73
US-10-719-956-525977
; Sequence 525977, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 525977
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-525977

Query Match 82.4%; Score 14; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCTTGCCGAG 17
Db 8 CACGCTTGCCGAG 21

RESULT 74
US-10-719-900-974704
; Sequence 974704, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 974704
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-974704

Query Match 82.4%; Score 14; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCG 15
Db 9 CCCACGCTTGCCG 22

RESULT 75
US-10-995-561-63916/c
; Sequence 63916, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF

```
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63916
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-63916

Query Match      82.4%; Score 14; DB 10; Length 201;
Best Local Similarity 87.5%; Pred. No. 2.1e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGCCCGAG 17
Db      101 MCCAGGCTTGCCCGAG 86
      :|||:|||||:|||||
      :|||:|||||:|||||

RESULT 76
US-10-301-480-242988/c
; Sequence 242988, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827,137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242988
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-242988

Query Match      82.4%; Score 14; DB 12; Length 429;
Best Local Similarity 87.5%; Pred. No. 1.9e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGCCCGAG 17
Db      280 CCCACGCTTGCCCGAG 265
      :|||:|||||:|||||
      :|||:|||||:|||||

RESULT 77
US-10-301-480-856397/c
; Sequence 856397, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827,137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 856397
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapien
```

```
US-10-301-480-856397

Query Match      82.4%; Score 14; DB 12; Length 429;
Best Local Similarity 87.5%; Pred. No. 1.9e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGCCCGAG 17
Db      280 CCCACGCTTGCCCGAG 265
      :|||:|||||:|||||
      :|||:|||||:|||||

RESULT 78
US-09-925-065A-148858/c
; Sequence 148858, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148858
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-148858

Query Match      82.4%; Score 14; DB 4; Length 440;
Best Local Similarity 87.5%; Pred. No. 1.9e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CCCACGCTTGCCCGAG 17
Db      279 CCCACGCTTGCCCGAG 264
      :|||:|||||:|||||
      :|||:|||||:|||||

RESULT 79
US-09-925-065A-148858/c
; Sequence 148858, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148858
```

;
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-148858

Query Match 82.4%; Score 14; DB 5; Length 440;
Best Local Similarity 87.5%; Pred. No. 1.9e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCACGCTGGCCGAG 17
|||:|||||
Db 279 CCCACGCTGGCCGAG 264

RESULT 80

US-09-918-995-28686/C
; Sequence 28686, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1998-01-20

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 28686
; LENGTH: 458
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-28686

Query Match 82.4%; Score 14; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTGGCCGCA 16
|||||:|||||
Db 358 CCACGCTGGCCGCA 345

RESULT 81

US-10-674-124A-19328
; Sequence 19328, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:

; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen

; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP

; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26

; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07

; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30

; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13

; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28

; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09

; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 19328

; LENGTH: 465

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AC073575.5_144468

; FEATURE:
; OTHER INFORMATION: Located on chromosome 12

; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on

; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base

; OTHER INFORMATION: sequence : 118577396

; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 99235

US-10-674-124A-19328

Query Match 82.4%; Score 14; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTGGCCGCA 16
|||||:|||||
Db 401 CCACGCTGGCCGCA 414

RESULT 82

US-09-925-065A-82285
; Sequence 82285, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single

; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135

; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 82285
; LENGTH: 518
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-925-065A-82285

Query Match 82.4%; Score 14; DB 4; Length 518;
Best Local Similarity 87.5%; Pred. No. 1.8e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACGCTGGCCGCA 16
|||||:|||||
Db 237 GCCCACGCTGGCCGCA 252

RESULT 83

US-09-925-065A-82285
; Sequence 82285, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single

; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135

; CURRENT APPLICATION NUMBER: US/09/925,065A

```
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82285
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-82285

Query Match      82.4%; Score 14; DB 5; Length 518;
Best Local Similarity 87.5%; Pred. No. 1.8e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  GCCCAGCGCTGGCCGA 16
        ||||| ||||| |||||
Db      237 GCCCAGCGCTGGCCRA 252

RESULT 84
US-10-301-480-183525
; Sequence 183525, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 183525
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-183525

Query Match      82.4%; Score 14; DB 12; Length 518;
Best Local Similarity 87.5%; Pred. No. 1.8e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  GCCCAGCGCTGGCCGA 16
        ||||| ||||| |||||
Db      237 GCCCAGCGCTGGCCRA 252

RESULT 85
US-10-301-480-796934
; Sequence 796934, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
```

```
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 796934
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-796934

Query Match      82.4%; Score 14; DB 12; Length 518;
Best Local Similarity 87.5%; Pred. No. 1.8e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  GCCCAGCGCTGGCCGA 16
        ||||| ||||| |||||
Db      237 GCCCAGCGCTGGCCRA 252

RESULT 86
US-09-925-065A-805681/c
; Sequence 805681, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805681
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-805681

Query Match      82.4%; Score 14; DB 4; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3  CCACGCTTGGCCGA 16
        ||||| ||||| |||||
Db      435 CCACGCTTGGCCGA 422

RESULT 87
US-09-925-065A-805681/c
; Sequence 805681, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
```



```

; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805681
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-805681

```

Query Match 82.4%; Score 14; DB 5; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTGGCCGA 16
Db 435 CCACGCTGGCCGA 422

```

RESULT 88
US-09-925-065A-14
; Sequence 14, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-14

```

Query Match 82.4%; Score 14; DB 4; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGCCGA 16
|||
Dp 99 CCACGCTTGCCGA 112

```

RESULT 89
US-09-925-065A-15
; Sequence 15, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827-135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24

```

```

, PRIOR APPLICATION NUMBER: US 60/252,147
,
, PRIOR FILING DATE: 2000-11-20
, PRIOR APPLICATION NUMBER: US 60/250,092
, PRIOR FILING DATE: 2000-11-30
, PRIOR APPLICATION NUMBER: US 60/261,766
, PRIOR FILING DATE: 2001-01-16
, PRIOR APPLICATION NUMBER: US 60/289,846
, PRIOR FILING DATE: 2001-05-09
, NUMBER OF SEQ ID NOS: 957086
, SOFTWARE: FastSEQ for Windows Version 4.0
, SEQ ID NO 15
, LENGTH: 683
, TYPE: DNA
, ORGANISM: Homo sapiens
US-09-925-065A-15

```

Query Match 82.4%; Score 14; DB 4; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGCCGA 16
Db 99 CCACGCTTGCCGA 112

```

RESULT 90
US-09-925-065A-14
; Sequence 14, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-14

```

```
Query Match      82.4%; Score 14; DB 5; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 3 CCACGCTGGCCGA 16
|||
Db 99 CCACGCTGGCCGA 112

```

RESULT 91
US-09-925-065A-15
; Sequence 15, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108027.135
; CURRENT APPLICATION NUMBER: US/09/925,065A

```

```
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-15

Query Match      82.4%; Score 14; DB 5; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGCCGA 16
   |||||
Db 99 CCACGCTTGCCGA 112

RESULT 92
US-10-301-480-101250
; Sequence 101250, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101250
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-101250

Query Match      82.4%; Score 14; DB 12; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGCCGA 16
   |||||
Db 99 CCACGCTTGCCGA 112

RESULT 93
US-10-301-480-101251
; Sequence 101251, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101251
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-101251

Query Match      82.4%; Score 14; DB 12; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGCCGA 16
   |||||
Db 99 CCACGCTTGCCGA 112

RESULT 94
US-10-301-480-714659
; Sequence 714659, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 714659
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-714659

Query Match      82.4%; Score 14; DB 12; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGCCGA 16
   |||||
Db 99 CCACGCTTGCCGA 112

RESULT 95
US-10-301-480-714660
; Sequence 714660, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 714660
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-714660
```

```
Query Match      82.4%; Score 14; DB 12; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGA 16
Db 99 CCACGCTTGGCCGA 112

RESULT 96
US-10-301-480-598304/c
; Sequence 598304, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 598304
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-598304

Query Match      82.4%; Score 14; DB 12; Length 911;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGA 16
Db 212 CCACGCTTGGCCGA 199

RESULT 97
US-10-301-480-1211713/c
; Sequence 1211713, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1211713
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1211713

Query Match      82.4%; Score 14; DB 12; Length 911;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCACGCTTGGCCGA 16
Db 212 CCACGCTTGGCCGA 199
```

```
RESULT 98
US-10-115-831-63/c
; Sequence 63, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyun
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 63
; LENGTH: 1181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)..(964)
US-10-115-831-63

Query Match      82.4%; Score 14; DB 7; Length 1181;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCG 15
Db 167 CCCACGCTTGGCCG 154

RESULT 99
US-11-096-568A-11881/c
; Sequence 11881, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 11881
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2475)
; OTHER INFORMATION: Ceres Seq. ID no. 13658778
US-11-096-568A-11881

Query Match      82.4%; Score 14; DB 16; Length 2475;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGGCC 14
Db 1894 GCCCACGCTTGGCC 1881

RESULT 100
US-10-995-561-13368/c
```

; Sequence 13368, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13368
; LENGTH: 70513
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13368

Query Match 82.4%; Score 14; DB 10; Length 70513;
Best Local Similarity 87.5%; Pred.No. 9.4e+02;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAG 17
Db 43202 MCCAGGCTTGGCCGAG 43187

Search completed: June 10, 2006, 20:52:22
Job time : 690.1 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:34:22 ; Search time 8.21667 Seconds
(without alignments)
262.600 Million cell updates/sec

Title: US-10-600-816-31

Perfect score: 17
Sequence: 1 gccacgcttgccgag 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 253354 seqs, 63461778 residues

Total number of hits satisfying chosen parameters: 506708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database : Published Applications NA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB_seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB_seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB_seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB_seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB_seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB_seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US12_NEW_PUB_seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US13_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	13.8	81.2	2165	6	US-10-511-937-485
2	13.8	81.2	5382	1	US-09-484-331-21
3	13.8	81.2	7873	6	US-10-485-397-7
4	13.8	81.2	7943	6	US-10-485-397-8
5	13.8	81.2	9737	1	US-09-484-331-22
6	13.8	81.2	9737	1	US-09-484-331-23
7	13.8	81.2	9737	1	US-09-484-331-28
8	13.8	81.2	9871	1	US-09-484-331-24
9	13.8	81.2	10060	1	US-09-484-331-25
C 10	13.4	78.8	806	6	US-10-953-349-15149
C 11	13.4	78.8	2019	6	US-10-953-349-35663
C 12	13.4	78.8	2175	7	US-11-293-697-559
C 13	13.4	78.8	2811	6	US-10-511-937-648
C 14	13.4	78.8	3048	7	US-11-293-697-950
15	12.8	75.3	39	7	US-11-169-140-80
16	12.8	75.3	39	7	US-11-169-140-84
C 17	12.8	75.3	39	7	US-11-319-606-12
C 18	12.8	75.3	500	7	US-11-249-305-1
C 19	12.8	75.3	500	7	US-11-249-305-2
C 20	12.8	75.3	500	7	US-11-249-305-3
C 21	12.8	75.3	500	7	US-11-249-305-4
C 22	12.8	75.3	693	7	US-11-217-529-82759
C 23	12.8	75.3	726	6	US-10-953-349-20941
C 24	12.8	75.3	777	7	US-11-217-529-82116
C 25	12.8	75.3	1086	7	US-11-217-529-1630

Sequence 25892, A	6	US-10-953-349-25892	1093	75.3	12.8	C 26
Sequence 33000, A	6	US-10-953-349-33000	1124	75.3	12.8	C 27
Sequence 35987, A	6	US-10-953-349-35987	1541	75.3	12.8	C 28
Sequence 33686, A	6	US-10-953-349-33686	1824	75.3	12.8	C 29
Sequence 1055, Ap	7	US-11-293-697-1055	2525	75.3	12.8	C 30
Sequence 1132, Ap	7	US-11-293-697-1132	2557	75.3	12.8	C 31
Sequence 845, App	7	US-11-293-697-845	2593	75.3	12.8	C 32
Sequence 1461, App	7	US-11-293-697-1461	2641	75.3	12.8	C 33
Sequence 1497, Ap	7	US-11-293-697-1497	2825	75.3	12.8	C 34
Sequence 891, App	7	US-11-293-697-885	3039	75.3	12.8	C 35
Sequence 291, App	7	US-11-293-697-291	3095	75.3	12.8	C 36
Sequence 1, Appli	7	US-11-254-185-1	3142	75.3	12.8	C 37
Sequence 510, App	7	US-11-293-697-510	3693	75.3	12.8	C 38
Sequence 158, App	7	US-11-121-154-158	5202	75.3	12.8	C 39
Sequence 1662, Ap	7	US-11-217-529-1662	9426	75.3	12.8	C 40
Sequence 568, App	6	US-10-505-928-568	10905	75.3	12.8	C 41
Sequence 34, Appl	6	US-10-857-260-34	27419	75.3	12.8	C 42
Sequence 151, App	6	US-10-505-928-151	128361	75.3	12.8	C 43
GENERAL INFORMATION	6	US-10-489-730-10	138941	75.3	12.8	C 44
Sequence 82, Appl	7	US-11-169-140-82	39	72.9	12.4	C 45
Sequence 87, Appl	7	US-11-169-140-87	39	72.9	12.4	C 46
Sequence 95, Appl	7	US-11-169-140-95	39	72.9	12.4	C 47
Sequence 26326, A	6	US-10-953-349-26326	506	72.9	12.4	C 48
Sequence 37335, A	6	US-10-953-349-37335	1035	72.9	12.4	C 49
Sequence 272, App	7	US-11-217-529-272	1242	72.9	12.4	C 50
Sequence 2011, Ap	6	US-10-471-571A-2011	1305	72.9	12.4	C 51
Sequence 35503, A	6	US-10-953-349-35503	1423	72.9	12.4	C 52
Sequence 35104, A	6	US-10-953-349-35104	1429	72.9	12.4	C 53
Sequence 3, Appli	7	US-11-293-697-3	1685	72.9	12.4	C 54
Sequence 559, App	7	US-11-293-697-559	2175	72.9	12.4	C 55
Sequence 322, App	7	US-11-293-697-322	2299	72.9	12.4	C 56
Sequence 1089, Ap	7	US-11-293-697-1089	2728	72.9	12.4	C 57
Sequence 1682, Ap	7	US-11-293-697-1682	2750	72.9	12.4	C 58
Sequence 576, App	6	US-10-511-937-576	2882	72.9	12.4	C 59
Sequence 3509, Ap	6	US-11-217-529-3509	3093	72.9	12.4	C 60
Sequence 5, Appli	6	US-10-527-552-5	16032	72.9	12.4	C 61
Sequence 64, Appl	7	US-11-189-279-64	118999	72.9	12.4	C 62
Sequence 11, Appl	7	US-11-259-950-11	522	71.8	12.2	C 63
Sequence 9, Appli	7	US-11-217-997-9	522	71.8	12.2	C 64
Sequence 35, Appl	7	US-11-217-997-35	522	71.8	12.2	C 65
Sequence 2, Appli	7	US-11-330-650-2	540	71.8	12.2	C 66
Sequence 1465, Ap	6	US-10-488-613-1465	562	71.8	12.2	C 67
Sequence 5, Appli	6	US-11-259-950-5	619	71.8	12.2	C 68
Sequence 3, Appli	6	US-10-953-349-14927	633	71.8	12.2	C 69
Sequence 2847, Ap	6	US-11-253-950-3	654	71.8	12.2	C 70
Sequence 566, App	6	US-10-488-619-2847	655	71.8	12.2	C 71
Sequence 1, Appli	7	US-11-217-529-566	681	71.8	12.2	C 72
Sequence 28486, A	6	US-11-259-950-1	762	71.8	12.2	C 73
Sequence 487, App	6	US-10-953-349-28486	798	71.8	12.2	C 74
Sequence 7, Appli	6	US-10-196-749-487	843	71.8	12.2	C 75
Sequence 20319, A	6	US-10-953-349-20319	868	71.8	12.2	C 76
Sequence 28667, A	6	US-11-217-997-7	877	71.8	12.2	C 77
Sequence 62, Appl	6	US-10-953-349-28667	945	71.8	12.2	C 78
Sequence 27232, A	6	US-11-256-428-62	949	71.8	12.2	C 79
Sequence 3387, Ap	6	US-10-953-349-27232	993	71.8	12.2	C 80
Sequence 35759, A	6	US-10-953-349-3387	1038	71.8	12.2	C 81
Sequence 1, Appli	6	US-11-254-252-1	1233	71.8	12.2	C 82
Sequence 77, Appl	7	US-11-259-950-77	1260	71.8	12.2	C 83
Sequence 9, Appli	7	US-11-259-950-9	1275	71.8	12.2	C 84
Sequence 27685, A	6	US-10-953-349-35397	1282	71.8	12.2	C 85
Sequence 28083, A	6	US-10-953-349-27685	1378	71.8	12.2	C 86
Sequence 35361, A	6	US-10-953-349-28083	1409	71.8	12.2	C 87
Sequence 653, App	6	US-10-505-928-653	1645	71.8	12.2	C 88
Sequence 981, App	6	US-11-293-697-981	1736	71.8	12.2	C 89
Sequence 34645, A	6	US-10-953-349-34645	1859	71.8	12.2	C 90
Sequence 13, Appl	6	US-11-254-252-13	1908	71.8	12.2	C 91
Sequence 79927, A	7	US-11-217-529-79927	1941	71.8	12.2	C 92
Sequence 2383, Ap	7	US-11-293-697-2383	2128	71.8	12.2	C 93
Sequence 1876, Ap	7	US-11-293-697-1876	2182	71.8	12.2	C 94

99	12.2	71.8	2208	7	US-11-293-697-40	Sequence 40, Appl	172	11.8	69.4	1047	6	US-10-471-571A-4381	Sequence 4381, Ap
c 100	12.2	71.8	2208	7	US-11-293-697-40	Sequence 40, Appl	c 173	11.8	69.4	1085	6	US-10-953-349-34244	Sequence 34244, A
c 101	12.2	71.8	2227	7	US-11-293-697-1054	Sequence 1054, Ap	174	11.8	69.4	1136	6	US-10-953-349-34021	Sequence 34021, A
c 102	12.2	71.8	2256	7	US-11-254-252-11	Sequence 11, Appl	c 175	11.8	69.4	1223	6	US-10-953-349-22839	Sequence 22839, A
c 103	12.2	71.8	2291	6	US-10-505-928-593	Sequence 593, App	c 176	11.8	69.4	1245	7	US-11-217-529-80547	Sequence 80547, A
c 104	12.2	71.8	2329	7	US-11-293-697-1489	Sequence 1489, Ap	c 177	11.8	69.4	1301	6	US-10-953-349-33917	Sequence 33917, A
c 105	12.2	71.8	2331	7	US-11-293-697-1950	Sequence 1950, Ap	c 178	11.8	69.4	1339	6	US-10-505-928-53	Sequence 53, Appl
c 106	12.2	71.8	2346	7	US-11-251-610-18	Sequence 18, Appl	c 179	11.8	69.4	1421	6	US-10-953-349-26978	Sequence 26978, A
c 107	12.2	71.8	2394	7	US-11-293-697-2296	Sequence 2296, Ap	c 180	11.8	69.4	1429	6	US-10-953-349-27636	Sequence 27636, A
c 108	12.2	71.8	2407	7	US-11-293-697-1675	Sequence 1602, Ap	c 181	11.8	69.4	1458	6	US-10-953-349-33142	Sequence 33142, A
c 109	12.2	71.8	2420	7	US-11-293-697-1602	Sequence 1675, Ap	c 182	11.8	69.4	1671	7	US-11-217-529-6137	Sequence 6137, Ap
c 110	12.2	71.8	2535	7	US-11-293-697-1400	Sequence 1400, Ap	c 183	11.8	69.4	1717	6	US-10-953-349-32286	Sequence 32286, A
c 111	12.2	71.8	2552	7	US-11-293-697-121	Sequence 121, App	c 184	11.8	69.4	1975	6	US-10-953-349-32079	Sequence 32079, A
c 112	12.2	71.8	2573	7	US-11-293-697-1384	Sequence 1384, Ap	c 185	11.8	69.4	2027	6	US-10-953-349-32588	Sequence 32588, A
c 113	12.2	71.8	2745	7	US-11-293-697-1028	Sequence 1028, Ap	c 186	11.8	69.4	2063	7	US-11-293-697-13	Sequence 13, Appl
c 114	12.2	71.8	2745	7	US-11-121-154-121	Sequence 121, App	c 187	11.8	69.4	2063	7	US-11-293-697-878	Sequence 878, App
c 115	12.2	71.8	2784	7	US-11-293-697-1348	Sequence 1348, Ap	c 188	11.8	69.4	2097	7	US-11-217-529-77826	Sequence 77826, A
c 116	12.2	71.8	2943	6	US-10-505-928-766	Sequence 766, App	c 189	11.8	69.4	2134	6	US-10-196-749-369	Sequence 369, App
c 117	12.2	71.8	3069	6	US-10-857-260-30	Sequence 30, Appl	c 190	11.8	69.4	2205	7	US-11-242-505A-23	Sequence 23, Appl
c 118	12.2	71.8	3417	7	US-11-131-901-7	Sequence 7, Appl	c 191	11.8	69.4	2291	7	US-11-293-697-1796	Sequence 1796, Ap
c 119	12.2	71.8	3445	7	US-11-293-697-899	Sequence 899, App	c 192	11.8	69.4	2312	7	US-11-293-697-218	Sequence 218, App
c 120	12.2	71.8	3646	6	US-10-511-937-358	Sequence 358, App	c 193	11.8	69.4	2388	6	US-10-196-749-218	Sequence 457, App
c 121	12.2	71.8	4255	7	US-11-217-937-37	Sequence 37, Appl	c 194	11.8	69.4	2439	7	US-11-217-529-5657	Sequence 5657, Ap
c 122	12.2	71.8	4733	7	US-11-217-937-5	Sequence 5, Appl	c 195	11.8	69.4	2452	7	US-11-242-505A-22	Sequence 22, Appl
c 123	12.2	71.8	4783	7	US-11-217-997-13	Sequence 13, Appl	c 196	11.8	69.4	2459	7	US-11-293-697-200	Sequence 200, App
c 124	12.2	71.8	4835	7	US-11-217-997-15	Sequence 15, Appl	c 197	11.8	69.4	2651	7	US-11-145-307A-73	Sequence 73, Appl
c 125	12.2	71.8	4896	7	US-11-258-767-61	Sequence 61, Appl	c 198	11.8	69.4	2724	7	US-11-217-529-5596	Sequence 5596, Ap
c 126	12.2	71.8	5000	7	US-11-217-997-19	Sequence 19, Appl	c 199	11.8	69.4	2733	6	US-10-511-937-577	Sequence 577, App
c 127	12.2	71.8	5005	7	US-11-217-997-39	Sequence 39, Appl	c 200	11.8	69.4	2802	7	US-11-121-154-180	Sequence 180, App
c 128	12.2	71.8	5005	7	US-11-217-997-21	Sequence 21, Appl	c 201	11.8	69.4	2805	7	US-11-145-307A-195	Sequence 195, App
c 129	12.2	71.8	5005	7	US-11-217-997-41	Sequence 41, Appl	c 202	11.8	69.4	2933	7	US-11-293-697-1383	Sequence 1383, Ap
c 130	12.2	71.8	5172	7	US-11-217-997-17	Sequence 17, Appl	c 203	11.8	69.4	2982	7	US-11-254-252-17	Sequence 17, Appl
c 131	12.2	71.8	5645	6	US-10-505-928-392	Sequence 392, App	c 204	11.8	69.4	3129	7	US-11-121-154-177	Sequence 177, App
c 132	12.2	71.8	5676	7	US-11-131-901-5	Sequence 5, Appl	c 205	11.8	69.4	3344	6	US-10-953-349-12071	Sequence 12071, A
c 133	12.2	71.8	6255	7	US-11-251-465-2	Sequence 2, Appl	c 206	11.8	69.4	3379	7	US-11-217-529-410	Sequence 410, App
c 134	12.2	71.8	7319	7	US-11-217-997-3	Sequence 3, Appl	c 207	11.8	69.4	4015	7	US-11-313-836-45	Sequence 45, Appl
c 135	12.2	71.8	7337	7	US-11-217-997-11	Sequence 11, Appl	c 208	11.8	69.4	5204	6	US-10-505-928-452	Sequence 452, App
c 136	12.2	71.8	7337	7	US-11-217-997-1	Sequence 1, Appl	c 209	11.8	69.4	7017	6	US-10-505-928-452	Sequence 452, App
c 137	12.2	71.8	9599	7	US-11-236-836-13	Sequence 13, Appl	c 210	11.8	69.4	11358	6	US-10-148-883-3	Sequence 3, Appl
c 138	12.2	71.8	9599	7	US-11-254-252-19	Sequence 19, Appl	c 211	11.8	69.4	12265	7	US-11-236-836-24	Sequence 24, Appl
c 139	12.2	71.8	12023	7	US-11-236-836-20	Sequence 20, Appl	c 212	11.8	69.4	13042	7	US-11-236-836-17	Sequence 17, Appl
c 140	12.2	71.8	12370	7	US-11-236-836-26	Sequence 26, Appl	c 213	11.8	69.4	13042	7	US-11-236-836-23	Sequence 23, Appl
c 141	12.2	71.8	12851	7	US-11-236-836-19	Sequence 19, Appl	c 214	11.8	69.4	15739	6	US-10-541-993-3	Sequence 3, Appl
c 142	12.2	71.8	13147	7	US-11-236-836-18	Sequence 18, Appl	c 215	11.8	69.4	16103	6	US-10-541-993-62	Sequence 62, Appl
c 143	12.2	71.8	13147	7	US-11-236-836-25	Sequence 25, Appl	c 216	11.8	69.4	16245	6	US-10-541-993-36	Sequence 36, Appl
c 144	12.2	71.8	27553	6	US-10-933-854-9	Sequence 9, Appl	c 217	11.8	69.4	16954	6	US-10-541-993-43	Sequence 43, Appl
c 145	12.2	71.8	38187	6	US-10-857-260-31	Sequence 31, Appl	c 218	11.8	69.4	16954	6	US-10-541-993-44	Sequence 44, Appl
c 146	12.2	71.8	38536	6	US-10-857-260-27	Sequence 27, Appl	c 219	11.8	69.4	17118	6	US-10-541-993-48	Sequence 48, Appl
c 147	12.2	71.8	42999	7	US-11-284-877-17	Sequence 17, Appl	c 220	11.8	69.4	17238	6	US-10-541-993-38	Sequence 38, Appl
c 148	12.2	71.8	138941	6	US-10-489-730-10	GENERAL INFORMATI	c 221	11.8	69.4	17238	6	US-10-541-993-39	Sequence 39, Appl
c 149	12.2	71.8	151830	6	US-10-519-335-37	Sequence 37, Appl	c 222	11.8	69.4	17593	6	US-10-541-993-42	Sequence 42, Appl
c 150	12	70.6	1114	6	US-10-953-349-30748	Sequence 30748, A	c 223	11.8	69.4	17593	6	US-10-541-993-47	Sequence 47, Appl
c 151	12	70.6	1272	6	US-10-953-349-7526	Sequence 7526, Ap	c 224	11.8	69.4	17877	6	US-10-541-993-37	Sequence 37, Appl
c 152	12	70.6	2158	7	US-11-245-628-9	Sequence 9, Appl	c 225	11.8	69.4	18333	6	US-10-541-993-51	Sequence 51, Appl
c 153	12	70.6	2408	7	US-11-293-697-225	Sequence 225, App	c 226	11.8	69.4	18449	6	US-10-541-993-40	Sequence 40, Appl
c 154	12	70.6	2817	7	US-11-293-697-892	Sequence 892, App	c 227	11.8	69.4	18449	6	US-10-541-993-41	Sequence 41, Appl
c 155	12	70.6	5515	6	US-10-524-021-1	Sequence 1, Appl	c 228	11.8	69.4	18449	6	US-10-541-993-49	Sequence 49, Appl
c 156	11.8	69.4	321	7	US-11-217-529-1920	Sequence 1920, Ap	c 229	11.8	69.4	18617	6	US-10-541-993-50	Sequence 50, Appl
c 157	11.8	69.4	396	6	US-10-560-723-156	Sequence 156, App	c 230	11.8	69.4	19491	6	US-10-541-993-45	Sequence 45, Appl
c 158	11.8	69.4	435	6	US-10-953-349-35785	Sequence 35785, A	c 231	11.8	69.4	21300	6	US-10-541-993-46	Sequence 46, Appl
c 159	11.8	69.4	435	6	US-10-488-619-608	Sequence 608, App	c 232	11.8	69.4	38187	6	US-10-857-260-31	Sequence 31, Appl
c 160	11.8	69.4	624	6	US-10-953-349-28259	Sequence 28259, A	c 233	11.8	69.4	42999	7	US-11-284-877-17	Sequence 17, Appl
c 161	11.8	69.4	819	7	US-11-217-529-47427	Sequence 17427, A	c 234	11.8	69.4	118899	7	US-11-189-279-64	Sequence 64, Appl
c 162	11.8	69.4	824	6	US-10-953-349-35520	Sequence 35520, A	c 235	11.4	67.1	25	7	US-11-217-529-149475	Sequence 149475,
c 163	11.8	69.4	836	6	US-10-953-349-34559	Sequence 34559, A	c 236	11.4	67.1	363	7	US-11-217-529-76617	Sequence 76617, A
c 164	11.8	69.4	838	6	US-10-953-349-40158	Sequence 40158, A	c 237	11.4	67.1	399	6	US-10-473-173-485	Sequence 485, App
c 165	11.8	69.4	873	6	US-10-953-349-25304	Sequence 25304, A	c 238	11.4	67.1	408	7	US-11-217-529-79607	Sequence 79607, A
c 166	11.8	69.4	924	7	US-11-217-529-4585	Sequence 4585, Ap	c 239	11.4	67.1	418	6	US-10-953-349-14494	Sequence 14494, A
c 167	11.8	69.4	960	7	US-11-121-154-78	Sequence 78, Appl	c 240	11.4	67.1	536	7	US-11-301-554-56	Sequence 56, Appl
c 168	11.8	69.4	999	6	US-10-471-571A-837	Sequence 837, App	c 241	11.4	67.1	616	6	US-10-488-619-1835	Sequence 1835, Ap
c 169	11.8	69.4	1000	7	US-11-233-726-13	Sequence 13, Appl	c 242	11.4	67.1	638	6	US-10-953-349-2907	Sequence 2907, Ap
c 170	11.8	69.4	1006	6	US-10-953-349-16158	Sequence 16158, A	c 243	11.4	67.1	705	6	US-10-953-349-35269	Sequence 35269, A
c 171	11.8	69.4	1024	6	US-10-196-749-943	Sequence 543, App	c 244	11.4	67.1	736	6	US-10-953-349-38808	Sequence 38808, A

C 245	11.4	67.1	780	7	US-11-217-529-80688	Sequence 80688, A	318	11.2	65.9	812	6	US-10-953-349-29397	Sequence 29397, A
C 246	11.4	67.1	887	7	US-10-953-349-39586	Sequence 39586, A	C 319	11.2	65.9	846	6	US-10-953-349-30869	Sequence 30869, A
C 247	11.4	67.1	897	7	US-11-217-529-2817	Sequence 2817, Ap	C 320	11.2	65.9	849	6	US-10-953-349-18027	Sequence 18027, A
C 248	11.4	67.1	943	6	US-10-953-349-30313	Sequence 30313, A	C 321	11.2	65.9	870	6	US-10-488-619-1839	Sequence 1839, Ap
C 249	11.4	67.1	1045	6	US-10-953-349-17086	Sequence 17086, A	C 322	11.2	65.9	870	6	US-11-154-103-30	Sequence 30, Appl
C 250	11.4	67.1	1159	6	US-10-953-349-25310	Sequence 25310, A	C 323	11.2	65.9	909	6	US-10-953-349-26362	Sequence 26362, A
C 251	11.4	67.1	1175	6	US-10-953-349-35242	Sequence 35242, A	C 324	11.2	65.9	909	6	US-10-953-349-35178	Sequence 35178, A
C 252	11.4	67.1	1176	6	US-10-953-349-29916	Sequence 29916, A	C 325	11.2	65.9	923	6	US-10-953-349-18891	Sequence 18891, A
C 253	11.4	67.1	1184	6	US-10-953-349-27363	Sequence 27363, A	C 326	11.2	65.9	944	6	US-10-953-349-40099	Sequence 40099, A
C 254	11.4	67.1	1260	7	US-11-256-428-69	Sequence 69, Appl	C 327	11.2	65.9	966	7	US-11-217-529-2506	Sequence 2506, Ap
C 255	11.4	67.1	1338	6	US-10-953-349-28670	Sequence 28670, A	C 328	11.2	65.9	978	7	US-11-217-529-695	Sequence 695, App
C 256	11.4	67.1	1424	6	US-10-953-349-15340	Sequence 15340, A	C 329	11.2	65.9	1008	7	US-11-217-529-1480	Sequence 1480, Ap
C 257	11.4	67.1	1515	7	US-11-217-529-4751	Sequence 4751, Ap	C 330	11.2	65.9	1041	7	US-11-121-154-131	Sequence 131, App
C 258	11.4	67.1	1524	7	US-11-217-529-5031	Sequence 5031, Ap	C 331	11.2	65.9	1054	6	US-10-468-193-21	Sequence 21, Appl
C 259	11.4	67.1	1674	7	US-11-249-111-18	Sequence 18, Appl	C 332	11.2	65.9	1072	6	US-10-953-349-39892	Sequence 39892, A
C 260	11.4	67.1	1704	6	US-10-953-349-30055	Sequence 30055, A	C 333	11.2	65.9	1089	7	US-11-217-529-80092	Sequence 80092, A
C 261	11.4	67.1	1754	7	US-11-293-697-1187	Sequence 1187, Ap	C 334	11.2	65.9	1114	6	US-10-953-349-39394	Sequence 39394, A
C 262	11.4	67.1	1854	7	US-11-293-697-3629	Sequence 3629, Ap	C 335	11.2	65.9	1136	6	US-10-953-349-22970	Sequence 22970, A
C 263	11.4	67.1	2003	7	US-11-293-697-11508	Sequence 11508, Ap	C 336	11.2	65.9	1143	7	US-11-217-529-80824	Sequence 80824, A
C 264	11.4	67.1	2106	7	US-11-217-529-81298	Sequence 81298, A	C 337	11.2	65.9	1150	6	US-10-953-349-1900	Sequence 1900, Ap
C 265	11.4	67.1	2154	7	US-11-293-697-137	Sequence 137, App	C 338	11.2	65.9	1161	7	US-11-297-134-9	Sequence 9, Appl
C 266	11.4	67.1	2235	7	US-11-293-697-935	Sequence 935, App	C 339	11.2	65.9	1165	6	US-10-953-349-36792	Sequence 36792, A
C 267	11.4	67.1	2379	6	US-10-196-749-483	Sequence 483, App	C 340	11.2	65.9	1209	7	US-11-264-737-51	Sequence 51, Appl
C 268	11.4	67.1	2440	7	US-11-293-697-569	Sequence 569, App	C 341	11.2	65.9	1209	7	US-11-265-761-39	Sequence 39, Appl
C 269	11.4	67.1	2525	7	US-11-293-697-1284	Sequence 1284, Ap	C 342	11.2	65.9	1212	6	US-10-953-349-24010	Sequence 24010, A
C 270	11.4	67.1	2528	7	US-11-293-697-1184	Sequence 1184, Ap	C 343	11.2	65.9	1226	6	US-10-953-349-36278	Sequence 36278, A
C 271	11.4	67.1	2623	7	US-11-293-697-1527	Sequence 1527, Ap	C 344	11.2	65.9	1254	6	US-10-953-349-34735	Sequence 34735, A
C 272	11.4	67.1	2643	6	US-10-953-349-38345	Sequence 38345, A	C 345	11.2	65.9	1299	6	US-10-953-349-11621	Sequence 11621, A
C 273	11.4	67.1	2681	7	US-11-293-697-1732	Sequence 1732, Ap	C 346	11.2	65.9	1305	7	US-11-217-529-174502	Sequence 174502, A
C 274	11.4	67.1	2763	7	US-11-293-697-1010	Sequence 1010, Ap	C 347	11.2	65.9	1320	6	US-10-370-959-65	Sequence 65, Appl
C 275	11.4	67.1	2769	7	US-11-293-697-280	Sequence 280, App	C 348	11.2	65.9	1321	6	US-10-196-749-371	Sequence 371, App
C 276	11.4	67.1	2809	7	US-11-293-697-2068	Sequence 2068, Ap	C 349	11.2	65.9	1326	7	US-11-217-529-1970	Sequence 1970, Ap
C 277	11.4	67.1	2958	7	US-11-217-529-44	Sequence 44, Appl	C 350	11.2	65.9	1360	6	US-10-953-349-32610	Sequence 32610, A
C 278	11.4	67.1	3094	7	US-11-293-697-306	Sequence 306, App	C 351	11.2	65.9	1391	6	US-10-953-349-28649	Sequence 28649, A
C 279	11.4	67.1	3098	7	US-11-293-697-1174	Sequence 1174, Ap	C 352	11.2	65.9	1470	6	US-10-953-349-27227	Sequence 27227, A
C 280	11.4	67.1	3308	7	US-11-293-697-1593	Sequence 1593, Ap	C 353	11.2	65.9	1507	6	US-10-953-349-34932	Sequence 34932, A
C 281	11.4	67.1	3399	7	US-11-217-529-73	Sequence 73, Appl	C 354	11.2	65.9	1516	6	US-10-953-349-17031	Sequence 17031, A
C 282	11.4	67.1	3417	7	US-11-293-697-2407	Sequence 2407, Ap	C 355	11.2	65.9	1585	6	US-10-953-349-25645	Sequence 25645, A
C 283	11.4	67.1	3487	7	US-11-293-697-419	Sequence 419, App	C 356	11.2	65.9	1587	7	US-11-293-697-1139	Sequence 1139, Ap
C 284	11.4	67.1	5658	7	US-11-217-529-249	Sequence 249, App	C 357	11.2	65.9	1602	6	US-10-953-349-32716	Sequence 32716, A
C 285	11.4	67.1	6263	6	US-10-560-723-142	Sequence 142, App	C 358	11.2	65.9	1617	6	US-10-196-749-235	Sequence 235, App
C 286	11.4	67.1	6600	6	US-10-525-318-5	Sequence 5, Appl	C 359	11.2	65.9	1617	7	US-11-101-316-61	Sequence 61, Appl
C 287	11.4	67.1	7000	6	US-10-525-318-8	Sequence 8, Appl	C 360	11.2	65.9	1681	6	US-10-953-349-36540	Sequence 36540, A
C 288	11.4	67.1	128361	6	US-10-505-928-151	Sequence 151, App	C 361	11.2	65.9	1700	7	US-11-158-743-1	Sequence 1, Appl
C 289	11.4	67.1	394191	6	US-10-506-549-3	Sequence 3, Appl	C 362	11.2	65.9	1705	6	US-10-953-349-3393	Sequence 393, App
C 290	11.2	65.9	25	7	US-11-217-529-13561	Sequence 13561, A	C 363	11.2	65.9	1733	6	US-10-953-349-11883	Sequence 11883, A
C 291	11.2	65.9	25	7	US-11-217-529-18256	Sequence 18256, A	C 364	11.2	65.9	1736	7	US-11-293-697-981	Sequence 981, App
C 292	11.2	65.9	25	7	US-11-217-529-11068	Sequence 11068, A	C 365	11.2	65.9	1767	6	US-10-953-349-22189	Sequence 22189, A
C 293	11.2	65.9	25	7	US-11-217-529-138713	Sequence 138713, A	C 366	11.2	65.9	1776	7	US-11-217-529-3257	Sequence 3257, Ap
C 294	11.2	65.9	32	7	US-11-291-668-33	Sequence 33, Appl	C 367	11.2	65.9	1781	6	US-10-953-349-6514	Sequence 6514, Ap
C 295	11.2	65.9	32	7	US-11-292-164-33	Sequence 61, Appl	C 368	11.2	65.9	1789	7	US-11-293-697-1140	Sequence 1140, Ap
C 296	11.2	65.9	39	7	US-11-156-014A-61	Sequence 61, Appl	C 369	11.2	65.9	1791	7	US-11-311-778-15	Sequence 15, Appl
C 297	11.2	65.9	50	6	US-10-554-711-575	Sequence 575, App	C 370	11.2	65.9	1804	7	US-11-251-465-7	Sequence 7, Appl
C 298	11.2	65.9	53	7	US-11-194-055-171	Sequence 171, App	C 371	11.2	65.9	1822	6	US-10-505-928-16	Sequence 16, Appl
C 299	11.2	65.9	89	7	US-11-194-055-170	Sequence 170, App	C 372	11.2	65.9	1827	1	US-09-949-925-13	Sequence 13, Appl
C 300	11.2	65.9	369	7	US-11-217-529-77870	Sequence 77870, A	C 373	11.2	65.9	1830	7	US-11-217-529-191102	Sequence 191102, A
C 301	11.2	65.9	369	7	US-11-217-529-173281	Sequence 173281, A	C 374	11.2	65.9	1831	6	US-10-953-349-11740	Sequence 11740, A
C 302	11.2	65.9	392	6	US-10-488-619-909	Sequence 909, App	C 375	11.2	65.9	1831	6	US-10-953-349-32436	Sequence 32436, A
C 303	11.2	65.9	486	6	US-10-488-619-443	Sequence 443, App	C 376	11.2	65.9	1836	6	US-10-953-349-35527	Sequence 35527, A
C 304	11.2	65.9	554	6	US-10-953-349-26889	Sequence 26889, A	C 377	11.2	65.9	1842	6	US-10-953-349-33234	Sequence 33234, A
C 305	11.2	65.9	634	6	US-10-953-349-25300	Sequence 25300, A	C 378	11.2	65.9	1849	7	US-10-953-349-25544	Sequence 25544, A
C 306	11.2	65.9	645	6	US-10-953-349-15352	Sequence 15352, A	C 379	11.2	65.9	1884	7	US-11-293-697-240	Sequence 240, App
C 307	11.2	65.9	663	7	US-11-217-529-2440	Sequence 2440, Ap	C 380	11.2	65.9	1900	6	US-10-953-349-4067	Sequence 4067, Ap
C 308	11.2	65.9	667	6	US-10-488-619-2188	Sequence 2188, Ap	C 381	11.2	65.9	1960	7	US-11-293-697-20	Sequence 20, Appl
C 309	11.2	65.9	669	6	US-10-953-349-26236	Sequence 26236, A	C 382	11.2	65.9	1967	7	US-11-293-697-962	Sequence 962, App
C 310	11.2	65.9	696	6	US-10-953-349-33816	Sequence 33816, A	C 383	11.2	65.9	1987	7	US-11-293-697-7240	Sequence 7240, Ap
C 311	11.2	65.9	717	6	US-10-953-349-25642	Sequence 25642, A	C 384	11.2	65.9	1999	6	US-10-953-349-134-64	Sequence 64, Appl
C 312	11.2	65.9	722	6	US-10-953-349-27119	Sequence 27119, A	C 385	11.2	65.9	2000	7	US-11-297-134-64	Sequence 64, Appl
C 313	11.2	65.9	732	7	US-11-217-529-1060	Sequence 1060, Ap	C 386	11.2	65.9	2026	7	US-11-293-697-1157	Sequence 1157, Ap
C 314	11.2	65.9	747	6	US-10-953-349-33906	Sequence 33906, A	C 387	11.2	65.9	2047	7	US-11-293-697-2016	Sequence 2016, Ap
C 315	11.2	65.9	773	6	US-10-953-349-37354	Sequence 37354, A	C 388	11.2	65.9	2058	6	US-10-370-959-63	Sequence 63, Appl
C 316	11.2	65.9	774	6	US-10-953-349-30246	Sequence 30246, A	C 389	11.2	65.9	2130	6	US-10-953-349-31754	Sequence 31754, A
C 317	11.2	65.9	789	7	US-11-217-529-5206	Sequence 5206, Ap	C 390	11.2	65.9	2139	7	US-11-245-628-15	Sequence 15, Appl

C 391	11.2	65.9	2143	6	US-10-953-349-38927	Sequence 38927, A	C 464	11	64.7	25	7	US-11-217-529-153221	Sequence 153221,
C 392	11.2	65.9	2175	6	US-10-953-349-25760	Sequence 25760, A	C 465	11	64.7	25	7	US-11-217-529-153240	Sequence 153240,
C 393	11.2	65.9	2214	6	US-10-953-349-34043	Sequence 34043, A	C 466	11	64.7	25	7	US-11-217-529-160103	Sequence 160103,
C 394	11.2	65.9	2280	7	US-11-217-529-80335	Sequence 80335, A	C 467	11	64.7	249	7	US-11-217-529-81634	Sequence 81634, A
C 395	11.2	65.9	2296	7	US-11-293-697-1115	Sequence 1, Appl	C 468	11	64.7	399	7	US-11-217-529-81634	Sequence 81634, A
C 396	11.2	65.9	2304	7	US-11-261-465-1	Sequence 1, Appl	C 469	11	64.7	438	6	US-10-488-619-1913	Sequence 1913, Ap
C 397	11.2	65.9	2306	7	US-11-293-697-2146	Sequence 2146, Ap	C 470	11	64.7	522	7	US-11-217-529-81632	Sequence 81632, A
C 398	11.2	65.9	2315	6	US-10-953-349-31973	Sequence 31973, A	C 471	11	64.7	647	7	US-11-321-421-338	Sequence 338, App
C 399	11.2	65.9	2405	7	US-11-293-697-1482	Sequence 1482, Ap	C 472	11	64.7	647	7	US-11-321-421-339	Sequence 339, App
C 400	11.2	65.9	2443	7	US-11-191-457-3	Sequence 3, Appl	C 473	11	64.7	705	7	US-11-217-529-1795	Sequence 1795, Ap
C 401	11.2	65.9	2476	6	US-10-196-749-585	Sequence 585, App	C 474	11	64.7	771	6	US-10-488-619-1914	Sequence 1914, Ap
C 402	11.2	65.9	2506	7	US-11-293-697-1843	Sequence 1843, Ap	C 475	11	64.7	822	6	US-10-953-349-33526	Sequence 33526, A
C 403	11.2	65.9	2510	7	US-11-293-697-1843	Sequence 1843, Ap	C 476	11	64.7	858	6	US-10-488-619-570	Sequence 570, App
C 404	11.2	65.9	2571	7	US-11-217-529-1119	Sequence 1119, Ap	C 477	11	64.7	870	7	US-11-217-529-82246	Sequence 82246, A
C 405	11.2	65.9	2590	7	US-11-293-697-238	Sequence 238, App	C 478	11	64.7	906	6	US-10-953-349-76132	Sequence 76132, A
C 406	11.2	65.9	2600	7	US-11-293-697-1576	Sequence 1576, Ap	C 479	11	64.7	1068	7	US-11-217-529-78106	Sequence 78106, A
C 407	11.2	65.9	2613	7	US-11-293-697-2367	Sequence 2367, Ap	C 480	11	64.7	1100	7	US-11-244-452-16	Sequence 16, Appl
C 408	11.2	65.9	2649	7	US-11-242-505A-46	Sequence 46, Appl	C 481	11	64.7	1167	7	US-11-113-081A-9	Sequence 9, Appl
C 409	11.2	65.9	2649	7	US-11-242-505A-47	Sequence 47, Appl	C 482	11	64.7	1203	7	US-11-024-544A-133	Sequence 133, App
C 410	11.2	65.9	2654	6	US-10-505-928-352	Sequence 352, App	C 483	11	64.7	1203	7	US-11-190-750-116	Sequence 116, App
C 411	11.2	65.9	2766	7	US-11-293-697-1119	Sequence 1119, Ap	C 484	11	64.7	1203	7	US-11-254-173-53	Sequence 53, Appl
C 412	11.2	65.9	2772	7	US-11-302-678-43	Sequence 43, Appl	C 485	11	64.7	1203	7	US-11-264-784-25	Sequence 25, Appl
C 413	11.2	65.9	2787	7	US-11-293-697-653	Sequence 653, App	C 486	11	64.7	1203	7	US-11-264-737-30	Sequence 30, Appl
C 414	11.2	65.9	2843	7	US-11-293-697-22	Sequence 22, Appl	C 487	11	64.7	1260	7	US-11-265-761-25	Sequence 25, Appl
C 415	11.2	65.9	2925	7	US-11-217-529-80109	Sequence 80109, A	C 488	11	64.7	1260	7	US-11-233-089-53	Sequence 53, Appl
C 416	11.2	65.9	2970	7	US-11-293-697-1467	Sequence 1467, Ap	C 489	11	64.7	1297	6	US-10-953-349-31295	Sequence 31295, A
C 417	11.2	65.9	2988	7	US-11-293-697-584	Sequence 584, App	C 490	11	64.7	1300	7	US-11-244-452-6	Sequence 6, Appl
C 418	11.2	65.9	3035	7	US-11-293-697-1791	Sequence 1791, Ap	C 491	11	64.7	1466	6	US-10-953-349-19994	Sequence 19994, A
C 419	11.2	65.9	3102	7	US-11-121-154-208	Sequence 208, App	C 492	11	64.7	1512	7	US-11-321-991-1	Sequence 1, Appl
C 420	11.2	65.9	3189	7	US-11-293-697-748	Sequence 748, App	C 493	11	64.7	1514	7	US-11-244-452-2	Sequence 2, Appl
C 421	11.2	65.9	3254	6	US-10-505-928-681	Sequence 681, App	C 494	11	64.7	1686	7	US-11-121-154-147	Sequence 147, App
C 422	11.2	65.9	3334	7	US-11-101-316-57	Sequence 57, Appl	C 495	11	64.7	1776	7	US-11-217-529-77895	Sequence 77895, A
C 423	11.2	65.9	3360	7	US-11-217-529-195613	Sequence 195613, A	C 496	11	64.7	1935	7	US-11-217-529-76590	Sequence 76590, A
C 424	11.2	65.9	3360	7	US-11-217-529-195615	Sequence 195615, A	C 497	11	64.7	1937	1	US-09-949-925-61	Sequence 61, Appl
C 425	11.2	65.9	3360	7	US-11-217-529-195615	Sequence 195615, A	C 498	11	64.7	1995	7	US-11-217-529-3311	Sequence 3311, Ap
C 426	11.2	65.9	3360	7	US-11-217-529-195637	Sequence 195637, A	C 499	11	64.7	2105	7	US-11-293-697-122	Sequence 122, App
C 427	11.2	65.9	3360	7	US-11-217-529-195638	Sequence 195638, A	C 500	11	64.7	2420	6	US-10-953-349-11350	Sequence 11350, A
C 428	11.2	65.9	3360	7	US-11-217-529-195639	Sequence 195639, A	C 501	11	64.7	2420	6	US-11-293-697-2342	Sequence 2342, Ap
C 429	11.2	65.9	3378	7	US-11-252-276-151	Sequence 151, App	C 502	11	64.7	2556	7	US-11-293-697-1531	Sequence 1531, Ap
C 430	11.2	65.9	3393	7	US-11-217-529-78969	Sequence 78969, A	C 503	11	64.7	2646	7	US-11-217-529-4274	Sequence 4274, Ap
C 431	11.2	65.9	3393	7	US-11-293-697-650	Sequence 650, App	C 504	11	64.7	3044	7	US-11-293-697-938	Sequence 938, App
C 432	11.2	65.9	3468	6	US-10-471-571A-4977	Sequence 4977, Ap	C 505	11	64.7	3210	7	US-11-217-529-4995	Sequence 4995, Ap
C 433	11.2	65.9	3489	7	US-11-106-014-93	Sequence 93, Appl	C 506	11	64.7	4628	6	US-10-511-937-2809	Sequence 2809, Ap
C 434	11.2	65.9	3530	6	US-11-293-697-521	Sequence 521, App	C 507	11	64.7	4628	6	US-10-511-937-2809	Sequence 2809, Ap
C 435	11.2	65.9	3608	6	US-10-196-749-433	Sequence 433, App	C 508	11	64.7	5094	6	US-10-505-928-701	Sequence 701, App
C 436	11.2	65.9	3633	7	US-11-217-529-110	Sequence 110, App	C 509	11	64.7	7425	7	US-11-217-529-77434	Sequence 77434, A
C 437	11.2	65.9	3931	7	US-11-293-697-532	Sequence 532, App	C 510	11	64.7	7425	7	US-11-254-173-49	Sequence 49, Appl
C 438	11.2	65.9	4004	7	US-11-293-697-590	Sequence 590, App	C 511	11	64.7	12663	7	US-11-264-784-132	Sequence 132, App
C 439	11.2	65.9	4277	6	US-10-196-749-439	Sequence 439, App	C 512	11	64.7	12663	7	US-11-264-737-160	Sequence 134, App
C 440	11.2	65.9	4833	7	US-11-252-276-21	Sequence 21, Appl	C 513	11	64.7	12663	7	US-11-265-761-134	Sequence 134, App
C 441	11.2	65.9	4932	6	US-10-511-937-379	Sequence 379, App	C 514	11	64.7	13321	7	US-11-264-784-136	Sequence 136, App
C 442	11.2	65.9	5022	7	US-11-217-529-3172	Sequence 3172, Ap	C 515	11	64.7	13321	7	US-11-264-737-188	Sequence 188, App
C 443	11.2	65.9	7686	6	US-10-511-937-652	Sequence 652, App	C 516	11	64.7	13321	7	US-11-265-761-140	Sequence 140, App
C 444	11.2	65.9	8194	7	US-11-024-544A-128	Sequence 128, App	C 517	11	64.7	22118	7	US-11-284-877-16	Sequence 16, Appl
C 445	11.2	65.9	8194	7	US-11-024-545-56	Sequence 56, Appl	C 518	11	64.7	33536	6	US-10-857-260-27	Sequence 27, Appl
C 446	11.2	65.9	8194	7	US-11-190-750-111	Sequence 111, App	C 519	10.8	63.5	25	7	US-11-217-529-6757	Sequence 6757, Ap
C 447	11.2	65.9	8194	7	US-11-251-466-30	Sequence 30, Appl	C 520	10.8	63.5	25	7	US-11-217-529-40096	Sequence 40096, A
C 448	11.2	65.9	8194	7	US-11-264-784-129	Sequence 129, App	C 521	10.8	63.5	25	7	US-11-217-529-73503	Sequence 73503, A
C 449	11.2	65.9	8194	7	US-11-264-737-183	Sequence 183, App	C 522	10.8	63.5	25	7	US-11-217-529-124332	Sequence 124332, A
C 450	11.2	65.9	8194	7	US-11-265-761-131	Sequence 131, App	C 523	10.8	63.5	25	7	US-11-217-529-160139	Sequence 160139, A
C 451	11.2	65.9	10189	7	US-11-264-737-176	Sequence 176, App	C 524	10.8	63.5	25	7	US-11-217-529-195285	Sequence 195285, A
C 452	11.2	65.9	10189	7	US-11-265-761-390	Sequence 390, App	C 525	10.8	63.5	39	7	US-11-169-140-88	Sequence 88, Appl
C 453	11.2	65.9	12558	7	US-11-264-737-177	Sequence 177, App	C 526	10.8	63.5	66	7	US-11-259-950-44	Sequence 44, Appl
C 454	11.2	65.9	12558	7	US-11-264-737-177	Sequence 177, App	C 527	10.8	63.5	250	7	US-11-136-524-103	Sequence 103, Appl
C 455	11.2	65.9	12558	7	US-11-265-761-391	Sequence 391, App	C 528	10.8	63.5	311	6	US-10-488-619-681	Sequence 681, App
C 456	11.2	65.9	12558	7	US-11-265-761-391	Sequence 391, App	C 529	10.8	63.5	312	7	US-11-217-529-6072	Sequence 6072, Ap
C 457	11.2	65.9	37426	6	US-10-473-173-32	Sequence 32, Appl	C 530	10.8	63.5	314	6	US-10-488-619-925	Sequence 925, App
C 458	11.2	65.9	37476	7	US-11-297-134-28	Sequence 28, Appl	C 531	10.8	63.5	374	6	US-10-488-619-1951	Sequence 1951, Ap
C 459	11.2	65.9	54550	7	US-11-318-813-42	Sequence 42, Appl	C 532	10.8	63.5	384	6	US-10-488-619-2568	Sequence 2568, Ap
C 460	11.2	65.9	135090	6	US-10-505-928-607	Sequence 607, App	C 533	10.8	63.5	390	6	US-10-488-619-882	Sequence 882, App
C 461	11.2	65.9	261789	7	US-11-260-842-1	Sequence 1, Appl	C 534	10.8	63.5	437	7	US-11-301-554-71	Sequence 71, Appl
C 462	11.2	65.9	394191	6	US-10-506-549-3	Sequence 3, Appl	C 535	10.8	63.5	441	6	US-10-560-723-9	Sequence 9, Appl
C 463	11.2	64.7	25	7	US-11-217-529-25765	Sequence 25765, A	C 536	10.8	63.5	453	7	US-11-217-529-4255	Sequence 4255, Ap

C 537	10.8	63.5	482	6	US-10-953-349-26070	Sequence 26070, A	610	10.8	63.5	1276	6	US-10-953-349-29067	Sequence 29067, A
C 538	10.8	63.5	516	6	US-10-953-349-25593	Sequence 25593, A	C 611	10.8	63.5	1281	7	US-11-217-529-79148	Sequence 79148, A
C 539	10.8	63.5	525	6	US-10-488-619-1331	Sequence 1331, A	C 612	10.8	63.5	1305	6	US-10-953-349-26630	Sequence 26630, A
C 540	10.8	63.5	531	6	US-10-488-619-2140	Sequence 2140, Ap	C 613	10.8	63.5	1344	7	US-11-217-529-6073	Sequence 6073, Ap
C 541	10.8	63.5	555	6	US-10-953-349-14804	Sequence 14804, A	C 614	10.8	63.5	1360	6	US-10-953-349-17739	Sequence 17739, A
C 542	10.8	63.5	557	1	US-09-949-925-75	Sequence 75, Appl	C 615	10.8	63.5	1375	6	US-10-953-349-29360	Sequence 29360, A
C 543	10.8	63.5	570	7	US-11-259-950-72	Sequence 72, Appl	C 616	10.8	63.5	1386	7	US-11-217-529-6153	Sequence 6153, Ap
C 544	10.8	63.5	571	6	US-10-488-619-2742	Sequence 2742, Ap	C 617	10.8	63.5	1386	6	US-10-953-349-21117	Sequence 21117, A
C 545	10.8	63.5	589	6	US-10-953-349-14023	Sequence 14023, A	C 618	10.8	63.5	1404	6	US-10-471-571A-449	Sequence 449, App
C 546	10.8	63.5	590	6	US-10-488-619-1759	Sequence 1759, Ap	C 619	10.8	63.5	1409	6	US-10-953-349-17924	Sequence 17924, A
C 547	10.8	63.5	597	7	US-11-259-950-74	Sequence 74, Appl	C 620	10.8	63.5	1409	6	US-10-953-349-18696	Sequence 18696, A
C 548	10.8	63.5	603	7	US-11-259-950-71	Sequence 71, Appl	C 621	10.8	63.5	1443	6	US-10-511-814-7	Sequence 7, Appli
C 549	10.8	63.5	603	7	US-11-259-950-73	Sequence 73, Appl	C 622	10.8	63.5	1461	6	US-10-953-349-32572	Sequence 32572, A
C 550	10.8	63.5	603	7	US-11-217-529-80	Sequence 80, Appl	C 623	10.8	63.5	1488	6	US-10-514-535-6	Sequence 6, Appli
C 551	10.8	63.5	608	6	US-10-953-349-38098	Sequence 80, Appl	C 624	10.8	63.5	1494	6	US-10-953-349-1037	Sequence 1037, Ap
C 552	10.8	63.5	630	7	US-11-217-529-131186	Sequence 38098, A	C 625	10.8	63.5	1500	6	US-10-505-928-210	Sequence 210, App
C 553	10.8	63.5	634	6	US-10-953-349-39182	Sequence 39182, A	C 626	10.8	63.5	1509	7	US-11-246-405-2	Sequence 2, Appli
C 554	10.8	63.5	636	7	US-11-259-950-88	Sequence 88, Appl	C 627	10.8	63.5	1511	6	US-10-505-928-90	Sequence 90, Appl
C 555	10.8	63.5	636	7	US-11-259-950-90	Sequence 90, Appl	C 628	10.8	63.5	1512	7	US-11-217-529-935	Sequence 935, App
C 556	10.8	63.5	636	7	US-11-259-950-92	Sequence 92, Appl	C 629	10.8	63.5	1542	6	US-10-953-349-30587	Sequence 30587, A
C 557	10.8	63.5	651	6	US-11-259-950-15	Sequence 15, Appl	C 630	10.8	63.5	1564	6	US-10-953-349-21513	Sequence 21513, A
C 558	10.8	63.5	660	6	US-10-196-749-467	Sequence 467, App	C 631	10.8	63.5	1569	7	US-11-217-529-78372	Sequence 78372, A
C 559	10.8	63.5	666	7	US-11-217-529-2735	Sequence 2735, Ap	C 632	10.8	63.5	1573	7	US-11-321-421-292	Sequence 292, App
C 560	10.8	63.5	721	6	US-10-953-349-25910	Sequence 25910, A	C 633	10.8	63.5	1573	7	US-11-321-421-294	Sequence 294, App
C 561	10.8	63.5	722	6	US-10-953-349-15388	Sequence 15388, A	C 634	10.8	63.5	1578	6	US-10-503-253A-3	Sequence 3, Appli
C 562	10.8	63.5	739	6	US-10-953-349-31113	Sequence 31113, A	C 635	10.8	63.5	1623	7	US-11-217-529-4591	Sequence 4591, Ap
C 563	10.8	63.5	754	6	US-10-953-349-30291	Sequence 30291, A	C 636	10.8	63.5	1631	7	US-11-293-697-1630	Sequence 1630, Ap
C 564	10.8	63.5	764	6	US-10-196-749-463	Sequence 463, App	C 637	10.8	63.5	1647	7	US-11-217-529-3089	Sequence 3089, Ap
C 565	10.8	63.5	799	6	US-10-488-619-838	Sequence 838, App	C 638	10.8	63.5	1656	6	US-11-293-697-1998	Sequence 1998, Ap
C 566	10.8	63.5	808	6	US-10-953-349-36860	Sequence 36860, A	C 639	10.8	63.5	1659	6	US-10-953-349-24585	Sequence 24585, A
C 567	10.8	63.5	825	7	US-11-217-529-222	Sequence 222, App	C 640	10.8	63.5	1660	6	US-10-953-349-21323	Sequence 21323, A
C 568	10.8	63.5	834	7	US-11-259-950-76	Sequence 76, Appl	C 641	10.8	63.5	1677	7	US-11-217-529-4477	Sequence 4477, Ap
C 569	10.8	63.5	865	6	US-10-953-349-8546	Sequence 8546, Ap	C 642	10.8	63.5	1704	7	US-11-217-529-5839	Sequence 5839, Ap
C 570	10.8	63.5	867	7	US-11-259-950-75	Sequence 75, Appl	C 643	10.8	63.5	1708	6	US-10-953-349-477	Sequence 477, App
C 571	10.8	63.5	870	7	US-11-217-529-76469	Sequence 76469, A	C 644	10.8	63.5	1710	7	US-11-217-529-1797	Sequence 1797, App
C 572	10.8	63.5	875	6	US-10-953-349-31120	Sequence 31120, A	C 645	10.8	63.5	1713	6	US-10-953-349-32050	Sequence 32050, A
C 573	10.8	63.5	883	7	US-11-256-428-48	Sequence 48, Appl	C 646	10.8	63.5	1713	7	US-11-217-529-2359	Sequence 2359, Ap
C 574	10.8	63.5	893	6	US-10-953-349-38142	Sequence 38142, A	C 647	10.8	63.5	1725	6	US-10-953-349-36039	Sequence 36039, A
C 575	10.8	63.5	901	6	US-10-953-349-17324	Sequence 17324, A	C 648	10.8	63.5	1750	6	US-10-953-349-20666	Sequence 20666, A
C 576	10.8	63.5	927	6	US-10-953-349-29997	Sequence 29997, A	C 649	10.8	63.5	1772	7	US-11-293-697-1877	Sequence 1877, Ap
C 577	10.8	63.5	927	6	US-10-953-349-29997	Sequence 29997, A	C 650	10.8	63.5	1779	6	US-11-217-529-82249	Sequence 82249, A
C 578	10.8	63.5	927	6	US-11-217-529-2379	Sequence 2379, Ap	C 651	10.8	63.5	1781	6	US-10-953-349-35542	Sequence 35542, A
C 579	10.8	63.5	934	6	US-10-953-349-33940	Sequence 33940, A	C 652	10.8	63.5	1783	7	US-11-293-697-2284	Sequence 2284, Ap
C 580	10.8	63.5	935	6	US-10-953-349-35334	Sequence 35334, A	C 653	10.8	63.5	1795	7	US-11-293-697-2063	Sequence 2063, Ap
C 581	10.8	63.5	953	6	US-10-953-349-29315	Sequence 29315, A	C 654	10.8	63.5	1808	7	US-11-293-697-382	Sequence 382, App
C 582	10.8	63.5	954	6	US-10-953-349-31361	Sequence 31361, A	C 655	10.8	63.5	1813	6	US-10-953-349-5321	Sequence 5321, Ap
C 583	10.8	63.5	967	6	US-10-953-349-31049	Sequence 31049, A	C 656	10.8	63.5	1856	7	US-11-293-697-893	Sequence 893, App
C 584	10.8	63.5	971	6	US-10-505-928-194	Sequence 194, App	C 657	10.8	63.5	1865	6	US-10-953-349-36070	Sequence 36070, A
C 585	10.8	63.5	984	7	US-11-217-529-749-85	Sequence 85, Appl	C 658	10.8	63.5	1885	6	US-10-953-349-25138	Sequence 25138, A
C 586	10.8	63.5	984	7	US-11-217-529-173918	Sequence 173918, A	C 659	10.8	63.5	1896	6	US-10-953-349-26191	Sequence 26191, A
C 587	10.8	63.5	993	7	US-11-249-111-16	Sequence 16, Appl	C 660	10.8	63.5	1896	7	US-11-293-697-81	Sequence 81, Appl
C 588	10.8	63.5	1000	6	US-10-953-349-38931	Sequence 38931, A	C 661	10.8	63.5	1914	7	US-11-293-697-2267	Sequence 2267, Ap
C 589	10.8	63.5	1086	7	US-11-217-529-2670	Sequence 2670, Ap	C 662	10.8	63.5	1920	7	US-11-293-697-1853	Sequence 1853, Ap
C 590	10.8	63.5	1086	7	US-11-217-529-2670	Sequence 2670, Ap	C 663	10.8	63.5	1923	7	US-11-145-307A-52	Sequence 52, Appl
C 591	10.8	63.5	1089	7	US-11-217-529-1497	Sequence 1496, Ap	C 664	10.8	63.5	1945	7	US-11-293-697-2153	Sequence 2153, Ap
C 592	10.8	63.5	1123	6	US-10-953-349-39069	Sequence 1497, Ap	C 665	10.8	63.5	1953	7	US-11-293-697-2125	Sequence 2125, Ap
C 593	10.8	63.5	1137	7	US-11-259-950-84	Sequence 84, Appl	C 666	10.8	63.5	1958	6	US-10-953-349-32934	Sequence 32934, A
C 594	10.8	63.5	1143	7	US-11-259-950-86	Sequence 86, Appl	C 667	10.8	63.5	1960	6	US-10-953-349-37443	Sequence 37443, A
C 595	10.8	63.5	1149	7	US-11-259-950-82	Sequence 82, Appl	C 668	10.8	63.5	1986	7	US-11-293-697-541	Sequence 541, App
C 596	10.8	63.5	1149	7	US-11-217-529-80114	Sequence 80114, A	C 669	10.8	63.5	1997	6	US-10-953-349-1486	Sequence 1486, Ap
C 597	10.8	63.5	1152	6	US-10-511-937-402	Sequence 402, App	C 670	10.8	63.5	1998	7	US-11-293-697-1063	Sequence 1063, Ap
C 598	10.8	63.5	1158	7	US-11-217-529-827	Sequence 827, App	C 671	10.8	63.5	1999	6	US-10-505-928-456	Sequence 456, App
C 599	10.8	63.5	1176	7	US-11-259-950-79	Sequence 79, Appl	C 672	10.8	63.5	2013	7	US-11-217-529-1571	Sequence 1571, Ap
C 600	10.8	63.5	1185	6	US-10-953-349-29401	Sequence 29401, A	C 673	10.8	63.5	2013	7	US-11-217-529-1527	Sequence 1527, Ap
C 601	10.8	63.5	1198	6	US-10-953-349-33186	Sequence 33186, A	C 674	10.8	63.5	2031	7	US-11-293-697-1582	Sequence 1582, Ap
C 602	10.8	63.5	1198	6	US-11-259-950-80	Sequence 80, Appl	C 675	10.8	63.5	2031	7	US-11-256-428-6	Sequence 6, Appli
C 603	10.8	63.5	1203	6	US-11-259-950-78	Sequence 78, Appl	C 676	10.8	63.5	2031	7	US-11-256-428-8	Sequence 8, Appli
C 604	10.8	63.5	1212	6	US-10-509-131-22	Sequence 22, Appl	C 677	10.8	63.5	2046	7	US-11-311-836-24	Sequence 24, Appl
C 605	10.8	63.5	1218	7	US-11-217-529-77474	Sequence 77474, A	C 678	10.8	63.5	2057	7	US-11-293-697-1042	Sequence 1042, Ap
C 606	10.8	63.5	1243	6	US-10-953-349-14254	Sequence 14254, A	C 679	10.8	63.5	2069	6	US-10-953-349-35949	Sequence 35949, A
C 607	10.8	63.5	1243	6	US-10-953-349-27577	Sequence 27577, A	C 680	10.8	63.5	2073	7	US-11-217-529-3340	Sequence 3340, Ap
C 608	10.8	63.5	1262	6	US-10-473-173-99	Sequence 99, Appl	C 681	10.8	63.5	2100	7	US-11-217-529-126	Sequence 126, App
C 609	10.8	63.5	1262	6	US-10-953-349-15668	Sequence 15668, A	C 682	10.8	63.5	2109	6	US-10-953-349-36514	Sequence 36514, A

c 683	10.8	63.5	2139	7	US-11-217-529-80809	Sequence 80809, A	c 756	10.8	63.5	2884	7	US-11-293-697-2232	Sequence 2232, Ap
c 684	10.8	63.5	2146	7	US-11-293-697-2119	Sequence 2119, Ap	c 757	10.8	63.5	2920	7	US-11-293-697-1041	Sequence 1041, Ap
c 685	10.8	63.5	2157	7	US-11-217-529-1378	Sequence 1378, Ap	c 758	10.8	63.5	2926	7	US-11-293-697-327	Sequence 327, App
c 686	10.8	63.5	2189	7	US-11-293-697-325	Sequence 325, App	c 759	10.8	63.5	2949	7	US-11-217-529-79063	Sequence 79063, A
c 687	10.8	63.5	2201	7	US-11-256-428-5	Sequence 5, Appli	c 760	10.8	63.5	2952	7	US-11-293-697-1728	Sequence 1728, Ap
c 688	10.8	63.5	2208	6	US-10-505-928-17	Sequence 17, Appli	c 761	10.8	63.5	2974	7	US-11-293-697-349	Sequence 349, App
c 689	10.8	63.5	2209	7	US-11-293-697-2220	Sequence 2220, Ap	c 762	10.8	63.5	2986	6	US-10-524-979-1	Sequence 1, Appli
c 690	10.8	63.5	2215	7	US-11-293-697-1751	Sequence 1751, Ap	c 763	10.8	63.5	3003	6	US-10-511-937-2820	Sequence 2820, Ap
c 691	10.8	63.5	2227	6	US-10-953-349-31181	Sequence 31181, A	c 764	10.8	63.5	3005	7	US-11-293-697-678	Sequence 678, App
c 692	10.8	63.5	2246	7	US-11-293-697-1694	Sequence 1694, Ap	c 765	10.8	63.5	3032	6	US-10-514-535-7	Sequence 7, Appli
c 693	10.8	63.5	2248	7	US-11-293-697-721	Sequence 721, App	c 766	10.8	63.5	3115	6	US-10-505-928-645	Sequence 645, App
c 694	10.8	63.5	2269	6	US-10-953-349-40059	Sequence 40059, A	c 767	10.8	63.5	3115	6	US-10-511-937-626	Sequence 626, App
c 695	10.8	63.5	2272	6	US-10-196-749-345	Sequence 345, App	c 768	10.8	63.5	3117	7	US-10-511-937-626	Sequence 626, App
c 696	10.8	63.5	2278	7	US-11-293-697-62	Sequence 62, Appli	c 769	10.8	63.5	3138	7	US-11-312-958-55	Sequence 55, Appli
c 697	10.8	63.5	2298	6	US-10-511-937-633	Sequence 633, App	c 770	10.8	63.5	3139	7	US-11-293-697-499	Sequence 499, App
c 698	10.8	63.5	2302	7	US-11-293-697-1520	Sequence 1520, Ap	c 771	10.8	63.5	3159	7	US-11-256-428-7	Sequence 7, Appli
c 699	10.8	63.5	2331	7	US-11-293-697-317	Sequence 317, App	c 772	10.8	63.5	3174	7	US-11-293-697-116	Sequence 116, App
c 700	10.8	63.5	2363	7	US-11-293-697-1688	Sequence 1688, Ap	c 773	10.8	63.5	3236	7	US-11-293-697-1225	Sequence 1225, Ap
c 701	10.8	63.5	2365	6	US-10-511-937-400	Sequence 400, App	c 774	10.8	63.5	3374	6	US-10-514-535-5	Sequence 5, Appli
c 702	10.8	63.5	2378	6	US-10-953-349-26450	Sequence 26450, A	c 775	10.8	63.5	3407	6	US-10-480-962-23	Sequence 23, Appli
c 703	10.8	63.5	2382	7	US-11-293-697-461	Sequence 461, App	c 776	10.8	63.5	3446	6	US-10-505-928-176	Sequence 176, App
c 704	10.8	63.5	2387	7	US-11-293-697-1487	Sequence 1487, Ap	c 777	10.8	63.5	3513	7	US-11-293-697-735	Sequence 735, App
c 705	10.8	63.5	2387	7	US-11-293-697-1941	Sequence 1941, Ap	c 778	10.8	63.5	3516	7	US-11-217-529-1430	Sequence 1430, Ap
c 706	10.8	63.5	2401	7	US-11-145-307A-258	Sequence 258, App	c 779	10.8	63.5	3663	7	US-11-217-529-1369	Sequence 1369, Ap
c 707	10.8	63.5	2402	7	US-11-293-697-261	Sequence 261, App	c 780	10.8	63.5	3668	7	US-11-293-697-1304	Sequence 1304, Ap
c 708	10.8	63.5	2402	7	US-11-293-697-1033	Sequence 1033, Ap	c 781	10.8	63.5	3671	6	US-10-196-749-265	Sequence 265, App
c 709	10.8	63.5	2405	7	US-11-293-697-1806	Sequence 1806, Ap	c 782	10.8	63.5	3680	6	US-10-196-749-173	Sequence 173, App
c 710	10.8	63.5	2414	7	US-11-293-697-350	Sequence 350, App	c 783	10.8	63.5	3733	6	US-10-514-535-3	Sequence 3, Appli
c 711	10.8	63.5	2419	7	US-11-293-697-1507	Sequence 1507, Ap	c 784	10.8	63.5	3733	6	US-10-514-535-3	Sequence 3, Appli
c 712	10.8	63.5	2442	7	US-11-217-529-80795	Sequence 80795, A	c 785	10.8	63.5	3780	6	US-10-509-131-24	Sequence 24, Appli
c 713	10.8	63.5	2454	6	US-10-511-937-2817	Sequence 2817, Ap	c 786	10.8	63.5	3780	6	US-10-509-131-24	Sequence 24, Appli
c 714	10.8	63.5	2468	7	US-11-293-697-1043	Sequence 1043, Ap	c 787	10.8	63.5	3787	6	US-10-514-535-4	Sequence 4, Appli
c 715	10.8	63.5	2484	7	US-11-293-697-2438	Sequence 2438, Ap	c 788	10.8	63.5	3787	6	US-10-514-535-4	Sequence 4, Appli
c 716	10.8	63.5	2506	7	US-11-293-697-54	Sequence 54, Appli	c 789	10.8	63.5	4226	7	US-11-293-697-1236	Sequence 1236, Ap
c 717	10.8	63.5	2507	7	US-11-293-697-400	Sequence 400, App	c 790	10.8	63.5	4226	6	US-10-505-928-544	Sequence 544, App
c 718	10.8	63.5	2519	6	US-10-953-349-36081	Sequence 36081, A	c 791	10.8	63.5	4246	6	US-10-505-928-544	Sequence 544, App
c 719	10.8	63.5	2522	6	US-10-504-120-7	Sequence 7, Appli	c 792	10.8	63.5	4326	7	US-11-121-154-160	Sequence 160, App
c 720	10.8	63.5	2535	7	US-11-293-697-2309	Sequence 2309, Ap	c 793	10.8	63.5	4326	7	US-11-248-986-28	Sequence 28, Appli
c 721	10.8	63.5	2541	7	US-11-293-697-2275	Sequence 2275, Ap	c 794	10.8	63.5	4326	7	US-11-248-986-28	Sequence 28, Appli
c 722	10.8	63.5	2574	7	US-11-293-697-674	Sequence 674, App	c 795	10.8	63.5	5132	6	US-10-505-928-663	Sequence 663, App
c 723	10.8	63.5	2575	7	US-11-293-697-736	Sequence 736, App	c 796	10.8	63.5	5132	6	US-10-505-928-663	Sequence 663, App
c 724	10.8	63.5	2592	7	US-11-293-697-444	Sequence 444, App	c 797	10.8	63.5	5132	6	US-10-505-928-663	Sequence 663, App
c 725	10.8	63.5	2596	7	US-11-293-697-385	Sequence 385, App	c 798	10.8	63.5	5132	6	US-10-505-928-663	Sequence 663, App
c 726	10.8	63.5	2622	7	US-11-217-529-79190	Sequence 79190, A	c 799	10.8	63.5	5132	6	US-10-505-928-663	Sequence 663, App
c 727	10.8	63.5	2642	7	US-11-293-697-2162	Sequence 2162, Ap	c 800	10.8	63.5	5132	6	US-10-505-928-663	Sequence 663, App
c 728	10.8	63.5	2649	6	US-10-505-928-736	Sequence 736, App	c 801	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 729	10.8	63.5	2649	7	US-11-242-505A-46	Sequence 46, Appli	c 802	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 730	10.8	63.5	2649	7	US-11-242-505A-47	Sequence 47, Appli	c 803	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 731	10.8	63.5	2652	7	US-11-293-697-351	Sequence 351, App	c 804	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 732	10.8	63.5	2655	7	US-11-293-697-1165	Sequence 1165, App	c 805	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 733	10.8	63.5	2676	7	US-11-217-529-79554	Sequence 79554, A	c 806	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 734	10.8	63.5	2682	7	US-11-293-697-1324	Sequence 1324, Ap	c 807	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 735	10.8	63.5	2684	6	US-10-509-131-52	Sequence 52, Appli	c 808	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 736	10.8	63.5	2688	7	US-11-293-697-2253	Sequence 2253, Ap	c 809	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 737	10.8	63.5	2696	7	US-11-317-329-2	Sequence 2, Appli	c 810	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 738	10.8	63.5	2696	7	US-11-317-329-4	Sequence 4, Appli	c 811	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 739	10.8	63.5	2696	7	US-11-317-329-5	Sequence 5, Appli	c 812	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 740	10.8	63.5	2696	7	US-11-317-329-6	Sequence 6, Appli	c 813	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 741	10.8	63.5	2696	7	US-11-317-329-7	Sequence 7, Appli	c 814	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 742	10.8	63.5	2696	7	US-11-317-329-8	Sequence 8, Appli	c 815	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 743	10.8	63.5	2696	7	US-11-317-329-9	Sequence 9, Appli	c 816	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 744	10.8	63.5	2696	7	US-11-317-329-10	Sequence 10, Appli	c 817	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 745	10.8	63.5	2741	7	US-11-326-489-4	Sequence 4, Appli	c 818	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 746	10.8	63.5	2752	6	US-10-953-349-33952	Sequence 33952, A	c 819	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 747	10.8	63.5	2752	7	US-11-293-697-598	Sequence 598, App	c 820	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 748	10.8	63.5	2754	6	US-10-953-349-3825	Sequence 3825, Ap	c 821	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 749	10.8	63.5	2777	7	US-11-293-697-578	Sequence 578, App	c 822	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 750	10.8	63.5	2821	6	US-10-505-928-734	Sequence 734, App	c 823	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 751	10.8	63.5	2823	7	US-11-293-697-2230	Sequence 2230, Ap	c 824	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 752	10.8	63.5	2824	7	US-11-293-697-1464	Sequence 1464, Ap	c 825	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 753	10.8	63.5	2842	7	US-11-293-697-1587	Sequence 1587, Ap	c 826	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 754	10.8	63.5	2854	6	US-10-473-173-64	Sequence 64, Appli	c 827	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App
c 755	10.8	63.5	2858	7	US-11-293-697-1363	Sequence 1363, Ap	c 828	10.8	63.5	6018	7	US-11-263-326-101	Sequence 101, App

C 829	10.8	63.5	26850	7	US-11-318-813-41	Sequence 41, Appl	C 902	10.6	62.4	750	6	US-10-953-349-16400	Sequence 16400, A
C 830	10.8	63.5	37426	6	US-10-473-173-32	Sequence 32, Appl	C 903	10.6	62.4	756	7	US-11-264-725-1	Sequence 1, Appl
C 831	10.8	63.5	37476	6	US-11-297-134-28	Sequence 28, Appl	C 904	10.6	62.4	760	6	US-10-953-349-28024	Sequence 28024, A
C 832	10.8	63.5	50000	6	US-10-528-659-2	Sequence 2, Appl	C 905	10.6	62.4	765	7	US-11-217-529-76391	Sequence 76391, A
C 833	10.8	63.5	50000	6	US-10-528-659-2	Sequence 2, Appl	C 906	10.6	62.4	767	7	US-11-266-747-2321	Sequence 2321, Ap
C 834	10.8	63.5	54550	7	US-11-318-813-42	Sequence 42, Appl	C 907	10.6	62.4	772	6	US-10-953-349-40073	Sequence 40073, A
C 835	10.8	63.5	55880	6	US-10-553-298-1	Sequence 1, Appl	C 908	10.6	62.4	783	6	US-10-505-928-297	Sequence 297, App
C 836	10.6	62.4	25	7	US-11-217-529-20027	Sequence 20027, A	C 909	10.6	62.4	783	6	US-10-953-349-9500	Sequence 9500, Ap
C 837	10.6	62.4	25	7	US-11-217-529-39285	Sequence 39285, A	C 910	10.6	62.4	783	7	US-11-217-529-800	Sequence 800, App
C 838	10.6	62.4	25	7	US-11-217-529-103436	Sequence 103436, A	C 911	10.6	62.4	783	7	US-11-090-563-23	Sequence 23, Appl
C 839	10.6	62.4	25	7	US-11-217-529-106911	Sequence 106911, A	C 912	10.6	62.4	795	6	US-10-488-619-2211	Sequence 2211, Ap
C 840	10.6	62.4	25	7	US-11-217-529-136975	Sequence 136975, A	C 913	10.6	62.4	810	6	US-10-953-349-34507	Sequence 34507, A
C 841	10.6	62.4	25	7	US-11-217-529-136986	Sequence 136986, A	C 914	10.6	62.4	832	6	US-10-953-349-34511	Sequence 34511, A
C 842	10.6	62.4	48	6	US-11-254-195-18	Sequence 18, Appl	C 915	10.6	62.4	839	6	US-10-953-349-28725	Sequence 28725, A
C 843	10.6	62.4	50	6	US-10-511-937-2676	Sequence 2676, Ap	C 916	10.6	62.4	843	6	US-10-488-619-1917	Sequence 1917, Ap
C 844	10.6	62.4	95	7	US-11-194-055-225	Sequence 225, App	C 917	10.6	62.4	851	6	US-10-488-619-2143	Sequence 2143, Ap
C 845	10.6	62.4	99	7	US-11-246-999-134	Sequence 134, App	C 918	10.6	62.4	858	6	US-10-953-349-28789	Sequence 28789, A
C 846	10.6	62.4	209	6	US-10-857-260-38	Sequence 38, Appl	C 919	10.6	62.4	862	6	US-10-953-349-33425	Sequence 33425, A
C 847	10.6	62.4	224	7	US-11-301-554-1613	Sequence 1613, Ap	C 920	10.6	62.4	886	6	US-10-953-349-34252	Sequence 34252, A
C 848	10.6	62.4	225	7	US-11-301-554-1770	Sequence 1770, Ap	C 921	10.6	62.4	886	6	US-10-953-349-38325	Sequence 38325, A
C 849	10.6	62.4	246	7	US-11-217-529-80184	Sequence 80184, A	C 922	10.6	62.4	901	6	US-10-953-349-3143	Sequence 3143, Ap
C 850	10.6	62.4	276	7	US-11-217-529-76043	Sequence 76043, A	C 923	10.6	62.4	903	6	US-10-953-349-39387	Sequence 39387, A
C 851	10.6	62.4	283	7	US-11-301-554-202	Sequence 202, App	C 924	10.6	62.4	925	6	US-10-953-349-28311	Sequence 28311, A
C 852	10.6	62.4	285	7	US-11-217-529-80183	Sequence 80183, A	C 925	10.6	62.4	937	6	US-10-953-349-25567	Sequence 25567, A
C 853	10.6	62.4	294	7	US-11-194-055-651	Sequence 651, App	C 926	10.6	62.4	956	6	US-10-953-349-34536	Sequence 34536, A
C 854	10.6	62.4	294	7	US-11-194-055-652	Sequence 652, App	C 927	10.6	62.4	958	6	US-10-953-349-39441	Sequence 39441, A
C 855	10.6	62.4	295	7	US-11-194-055-653	Sequence 653, App	C 928	10.6	62.4	977	6	US-10-953-349-30709	Sequence 30709, A
C 856	10.6	62.4	300	6	US-10-488-619-2178	Sequence 2178, Ap	C 929	10.6	62.4	981	6	US-10-953-349-30923	Sequence 30923, A
C 857	10.6	62.4	305	6	US-10-560-723-130	Sequence 130, App	C 930	10.6	62.4	981	6	US-10-953-349-30923	Sequence 30923, A
C 858	10.6	62.4	315	6	US-10-488-619-203	Sequence 203, App	C 931	10.6	62.4	999	7	US-11-217-529-76891	Sequence 76891, A
C 859	10.6	62.4	366	7	US-11-217-529-77191	Sequence 77191, A	C 932	10.6	62.4	1006	6	US-10-953-349-16158	Sequence 16158, A
C 860	10.6	62.4	379	6	US-10-488-619-780	Sequence 780, App	C 933	10.6	62.4	1009	6	US-10-953-349-27931	Sequence 27931, A
C 861	10.6	62.4	427	6	US-10-488-619-1587	Sequence 1587, Ap	C 934	10.6	62.4	1012	6	US-10-953-349-16083	Sequence 16083, A
C 862	10.6	62.4	450	7	US-11-217-529-166997	Sequence 166997, A	C 935	10.6	62.4	1024	6	US-10-953-349-5949	Sequence 5949, Ap
C 863	10.6	62.4	455	6	US-10-488-619-724	Sequence 724, App	C 936	10.6	62.4	1044	7	US-11-217-529-1278	Sequence 1278, Ap
C 864	10.6	62.4	463	6	US-10-488-619-360	Sequence 360, App	C 937	10.6	62.4	1053	6	US-10-953-349-34169	Sequence 34169, A
C 865	10.6	62.4	468	6	US-10-953-349-30051	Sequence 30051, A	C 938	10.6	62.4	1053	7	US-11-217-529-6204	Sequence 6204, Ap
C 866	10.6	62.4	472	7	US-11-301-554-835	Sequence 835, App	C 939	10.6	62.4	1056	7	US-11-249-111-50	Sequence 50, Appl
C 867	10.6	62.4	473	7	US-11-170-482-4	Sequence 4, Appl	C 940	10.6	62.4	1062	7	US-11-217-529-1077	Sequence 1077, Ap
C 868	10.6	62.4	478	7	US-11-145-307A-9	Sequence 9, Appl	C 941	10.6	62.4	1086	7	US-11-253-300-40	Sequence 40, Appl
C 869	10.6	62.4	481	6	US-10-488-619-2572	Sequence 2572, Ap	C 942	10.6	62.4	1089	7	US-11-217-529-3016	Sequence 3016, Ap
C 870	10.6	62.4	488	6	US-10-953-349-2768	Sequence 2768, Ap	C 943	10.6	62.4	1097	6	US-10-953-349-33019	Sequence 33019, A
C 871	10.6	62.4	507	7	US-11-122-986-211	Sequence 211, App	C 944	10.6	62.4	1101	7	US-11-253-300-28	Sequence 28, Appl
C 872	10.6	62.4	507	7	US-11-122-986-211	Sequence 213, App	C 945	10.6	62.4	1135	6	US-10-953-349-34319	Sequence 34319, A
C 873	10.6	62.4	515	6	US-10-488-619-902	Sequence 902, App	C 946	10.6	62.4	1150	6	US-10-953-349-17320	Sequence 17320, A
C 874	10.6	62.4	517	6	US-10-953-349-33449	Sequence 33449, A	C 947	10.6	62.4	1151	6	US-10-953-349-31677	Sequence 31677, A
C 875	10.6	62.4	528	7	US-11-217-529-173659	Sequence 173659, A	C 948	10.6	62.4	1156	6	US-10-953-349-31118	Sequence 31118, A
C 876	10.6	62.4	529	6	US-10-953-349-30756	Sequence 30756, A	C 949	10.6	62.4	1163	6	US-10-953-349-11895	Sequence 11895, A
C 877	10.6	62.4	539	6	US-10-488-619-2573	Sequence 2573, Ap	C 950	10.6	62.4	1184	6	US-10-953-349-20370	Sequence 20370, A
C 878	10.6	62.4	543	7	US-11-217-529-5827	Sequence 5827, Ap	C 951	10.6	62.4	1197	6	US-10-953-349-10496	Sequence 10496, A
C 879	10.6	62.4	548	6	US-10-488-619-2823	Sequence 2823, Ap	C 952	10.6	62.4	1198	6	US-10-953-349-22337	Sequence 22337, A
C 880	10.6	62.4	556	6	US-10-488-619-2813	Sequence 2813, Ap	C 953	10.6	62.4	1205	6	US-10-953-349-28073	Sequence 28073, A
C 881	10.6	62.4	562	6	US-10-953-349-35876	Sequence 35876, A	C 954	10.6	62.4	1207	6	US-10-953-349-32165	Sequence 32165, A
C 882	10.6	62.4	566	6	US-10-488-619-2661	Sequence 2661, Ap	C 955	10.6	62.4	1211	6	US-10-953-349-32165	Sequence 32165, A
C 883	10.6	62.4	569	6	US-10-488-619-2324	Sequence 2324, Ap	C 956	10.6	62.4	1227	6	US-10-953-349-35524	Sequence 35524, A
C 884	10.6	62.4	571	6	US-10-953-349-35685	Sequence 35685, A	C 957	10.6	62.4	1269	6	US-10-474-853-6	Sequence 6, Appl
C 885	10.6	62.4	575	6	US-10-511-937-2890	Sequence 2890, Ap	C 958	10.6	62.4	1280	6	US-10-953-349-22574	Sequence 22574, A
C 886	10.6	62.4	582	6	US-10-953-349-29867	Sequence 29867, A	C 959	10.6	62.4	1284	7	US-11-217-529-76846	Sequence 76846, A
C 887	10.6	62.4	584	6	US-10-488-619-2964	Sequence 2964, Ap	C 960	10.6	62.4	1302	6	US-10-953-349-33793	Sequence 33793, A
C 888	10.6	62.4	595	6	US-10-488-619-2212	Sequence 2212, Ap	C 961	10.6	62.4	1305	7	US-11-217-529-2401	Sequence 2401, Ap
C 889	10.6	62.4	596	7	US-11-156-014A-107	Sequence 107, App	C 962	10.6	62.4	1340	6	US-10-953-349-23384	Sequence 23384, A
C 890	10.6	62.4	609	6	US-10-488-619-3022	Sequence 3022, Ap	C 963	10.6	62.4	1349	6	US-10-953-349-34806	Sequence 34806, A
C 891	10.6	62.4	625	6	US-10-953-349-29539	Sequence 29539, A	C 964	10.6	62.4	1354	6	US-10-953-349-27589	Sequence 27589, A
C 892	10.6	62.4	655	6	US-10-488-619-2847	Sequence 2847, Ap	C 965	10.6	62.4	1371	7	US-11-217-529-173711	Sequence 173711, A
C 893	10.6	62.4	684	6	US-10-953-349-21645	Sequence 21645, A	C 966	10.6	62.4	1374	7	US-11-264-737-1	Sequence 1, Appl
C 894	10.6	62.4	685	6	US-10-953-349-38375	Sequence 38375, A	C 967	10.6	62.4	1374	7	US-11-265-761-1	Sequence 1, Appl
C 895	10.6	62.4	699	6	US-10-953-349-16591	Sequence 16591, A	C 968	10.6	62.4	1376	6	US-10-953-349-28530	Sequence 28530, A
C 896	10.6	62.4	706	6	US-10-953-349-31527	Sequence 31527, A	C 969	10.6	62.4	1376	6	US-10-953-349-34157	Sequence 34157, A
C 897	10.6	62.4	712	6	US-10-488-619-2963	Sequence 2963, Ap	C 970	10.6	62.4	1376	6	US-10-953-349-195-3	Sequence 3, Appl
C 898	10.6	62.4	723	6	US-10-473-173-255	Sequence 255, App	C 971	10.6	62.4	1382	7	US-11-254-195-3	Sequence 3, Appl
C 899	10.6	62.4	723	6	US-10-953-349-12320	Sequence 12320, A	C 972	10.6	62.4	1382	7	US-11-217-529-80786	Sequence 80786, A
C 900	10.6	62.4	733	6	US-10-953-349-32724	Sequence 32724, A	C 973	10.6	62.4	1401	6	US-10-953-349-27097	Sequence 27097, A
C 901	10.6	62.4	738	6	US-10-511-937-364	Sequence 364, App	C 974	10.6	62.4	1401	6	US-10-953-349-27097	Sequence 27097, A

```
c 975 10.6 62.4 1408 6 US-10-953-349-30818
c 976 10.6 62.4 1411 6 US-10-953-349-38703
c 977 10.6 62.4 1437 7 US-11-217-529-78974
c 978 10.6 62.4 1439 6 US-10-953-349-32310
c 979 10.6 62.4 1446 6 US-10-953-349-35850
c 980 10.6 62.4 1447 7 US-11-243-296A-4
c 981 10.6 62.4 1449 7 US-11-217-529-2200
c 982 10.6 62.4 1451 6 US-10-953-349-31780
c 983 10.6 62.4 1464 6 US-10-953-349-29604
c 984 10.6 62.4 1466 6 US-10-505-928-458
c 985 10.6 62.4 1466 6 US-10-505-928-458
c 986 10.6 62.4 1485 6 US-10-953-349-24486
c 987 10.6 62.4 1486 6 US-10-953-349-22309
c 988 10.6 62.4 1486 6 US-10-953-349-38908
c 989 10.6 62.4 1504 6 US-10-953-349-35108
c 990 10.6 62.4 1517 6 US-10-953-349-35879
c 991 10.6 62.4 1529 6 US-10-953-349-33663
c 992 10.6 62.4 1531 6 US-10-953-349-34579
c 993 10.6 62.4 1533 6 US-10-953-349-21987
c 994 10.6 62.4 1541 6 US-10-953-349-12807
c 995 10.6 62.4 1555 6 US-10-953-349-34789
c 996 10.6 62.4 1556 6 US-10-953-349-30645
c 997 10.6 62.4 1568 6 US-10-953-349-24028
c 998 10.6 62.4 1571 6 US-10-713-648A-39
c 999 10.6 62.4 1576 6 US-10-953-349-7463
c1000 10.6 62.4 1579 7 US-11-293-697-1142
```

ALIGNMENTS

```
RESULT 1
US-10-511-937-485/c
; Sequence 485, Application US/10511937
; Publication No. US20060089836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 485
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-485
```

```
Query Match 81.2%; Score 13.8; DB 6; Length 2165;
Best Local Similarity 88.2%; Pred. No. 39;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GCCCAGCTTGCCGAG 17
||| ||||| ||||| |||||
Db 262 GCCCAGCTTGCCGAG 246
```

RESULT 2

```
US-09-484-331-21
; Sequence 21, Application US/09484331
; Publication No. US20060105318A1
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; TITLE OF INVENTION: ENDOGENOUS GENES
; FILE REFERENCE: 0221-0003L
; CURRENT APPLICATION NUMBER: US/09/484,331
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 5382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (890)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1042)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-331-21
```

```
Query Match 81.2%; Score 13.8; DB 1; Length 5382;
Best Local Similarity 88.2%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GCCCAGCTTGCCGAG 17
||| ||||| ||||| |||||
Db 1501 GCCCAGCTTGCCGAG 1517
```

RESULT 3

```
US-10-485-397-7
; Sequence 7, Application US/10485397
; Publication No. US20060099673A1
; GENERAL INFORMATION:
; APPLICANT: ALTANA Pharma AG
; TITLE OF INVENTION: Novel Recombinant Gene Expression Method
; FILE REFERENCE: B697USPCT01
; CURRENT APPLICATION NUMBER: US/10/485,397
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 7873
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: pGFPstopneo is a circular Plasmid DNA
US-10-485-397-7
```

```
Query Match 81.2%; Score 13.8; DB 6; Length 7873;
Best Local Similarity 88.2%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GCCCAGCTTGCCGAG 17
||| ||||| ||||| |||||
```

```
Db      5129 GCCCGCGCATGGCCGAG 5145

RESULT 4
US-10-485-397-8
; Sequence 8, Application US/10485397
; Publication No. US20060099673A1
; GENERAL INFORMATION:
; APPLICANT: ALTANA Pharma AG
; TITLE OF INVENTION: Novel Recombinant Gene Expression Method
; FILE REFERENCE: B697USCT01
; CURRENT APPLICATION NUMBER: US/10/485,397
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 7943
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: pGFPstopneosecis is a circular Plasmid DNA
US-10-485-397-8

Query Match      81.2%; Score 13.8; DB 6; Length 7943;
Best Local Similarity 88.2%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGGCCGAG 17
      ||||| ||| ||||| |||||
Db      5199 GCCCGCGCATGGCCGAG 5215

RESULT 5
US-09-484-331-22
; Sequence 22, Application US/09484331
; Publication No. US20060105318A1
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003L
; CURRENT APPLICATION NUMBER: US/09/484,331
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-331-22

Query Match      81.2%; Score 13.8; DB 1; Length 9737;
Best Local Similarity 88.2%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGGCCGAG 17
      ||||| ||| ||||| |||||
Db      8958 GCCCGCGCATGGCCGAG 8974

RESULT 6
US-09-484-331-23
; Sequence 23, Application US/09484331
; Publication No. US20060105318A1
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003L
; CURRENT APPLICATION NUMBER: US/09/484,331
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-331-23

Query Match      81.2%; Score 13.8; DB 1; Length 9737;
Best Local Similarity 88.2%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GCCCAGCTTGGCCGAG 17
      ||||| ||| ||||| |||||
Db      8958 GCCCGCGCATGGCCGAG 8974

RESULT 7
US-09-484-331-28
; Sequence 28, Application US/09484331
; Publication No. US20060105318A1
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003L
; CURRENT APPLICATION NUMBER: US/09/484,331
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/253,022
```



```

; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/159,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8347)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8499)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-331-28

Query Match      81.2%; Score 13.8; DB 1; Length 9737;
Best Local Similarity 88.2%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGTTGGCCGAG 17
Db      8958 GCCCGCGCATGGCCGAG 8974

RESULT 8
US-09-484-331-24
; Sequence 24, Application US/09484331
; Publication No. US20060105318A1
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003L
; CURRENT APPLICATION NUMBER: US/09/484,331
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8481)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8633)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-331-24

Query Match      81.2%; Score 13.8; DB 1; Length 9871;
Best Local Similarity 88.2%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGTTGGCCGAG 17
Db      8958 GCCCGCGCATGGCCGAG 8974

RESULT 9
US-09-484-331-25
; Sequence 25, Application US/09484331
; Publication No. US20060105318A1
; GENERAL INFORMATION:
; APPLICANT: HARRINGTON, JOHN J.
; APPLICANT: SHERF, BRUCE
; APPLICANT: RUNDLETT, STEPHEN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NON-TARGETED ACTIVATION OF
; FILE REFERENCE: 0221-0003L
; CURRENT APPLICATION NUMBER: US/09/484,331
; CURRENT FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/276,820
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 09/263,814
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/253,022
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 08/941,223
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 10060
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8670)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (8822)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-484-331-25

Query Match      81.2%; Score 13.8; DB 1; Length 10060;
Best Local Similarity 88.2%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  GCCCAGCGTTGGCCGAG 17
Db      9281 GCCCGCGCATGGCCGAG 9297

RESULT 10
US-10-953-349-15149/c
; Sequence 15149, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15149
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-15149

Query Match      78.8%; Score 13.4; DB 6; Length 806;

```



```
Best Local Similarity 93.3%; Pred. No. 61;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGAG 17
DB 492 CCATGCTTGGCCGAG 478

RESULT 11
US-10-953-349-35663/c
; Sequence 35663, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35663
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-35663

Query Match 78.8%; Score 13.4; DB 6; Length 2019;
Best Local Similarity 93.3%; Pred. No. 64;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGAG 17
DB 203 CCACGCTTGGCCGCG 189

RESULT 12
US-11-293-697-559
; Sequence 559, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 559
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-559

Query Match 78.8%; Score 13.4; DB 7; Length 2175;
Best Local Similarity 93.3%; Pred. No. 64;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGAG 17
DB 1948 CCACGCTGGCCGAG 1962

RESULT 13
US-10-511-937-648/c
; Sequence 648, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
```

```
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 648
; LENGTH: 2811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-648

Query Match 78.8%; Score 13.4; DB 6; Length 2811;
Best Local Similarity 93.3%; Pred. No. 65;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCCGAG 17
DB 1465 CCACCTCTTGGCCGAG 1451

RESULT 14
US-11-293-697-950
; Sequence 950, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 950
; LENGTH: 3048
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-950

Query Match 78.8%; Score 13.4; DB 7; Length 3048;
Best Local Similarity 93.3%; Pred. No. 65;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGGCCGA 16
DB 838 CCCACGCTTGGCCAA 852

RESULT 15
US-11-169-140-80
; Sequence 80, Application US/11169140
; Publication No. US20060099150A1
; GENERAL INFORMATION:
; APPLICANT: ARIZKE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven
```



```
RESULT 19
US-11-249-305-2
; Sequence 2, Application US/11249305
; Publication No. US20060115836A1
; GENERAL INFORMATION:
; APPLICANT: TRINQUET, VALERIE
; APPLICANT: ROSSIO, PATRICIA
; TITLE OF INVENTION: THE ULTIMATE TREATMENT OF PSORIASIS
; FILE REFERENCE: 034227-098
; CURRENT APPLICATION NUMBER: US/11/249,305
; CURRENT FILING DATE: 2005-10-14
; PRIOR APPLICATION NUMBER: PCT/FR04/000934
; PRIOR FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: FR 0304773
; PRIOR FILING DATE: 2003-04-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-249-305-2

Query Match 75.3%; Score 12.8; DB 7; Length 500;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCCGA 16
Db 424 GCCCAGCGCTTGGCCGA 439

RESULT 20
US-11-249-305-3/c
; Sequence 3, Application US/11249305
; Publication No. US20060115836A1
; GENERAL INFORMATION:
; APPLICANT: TRINQUET, VALERIE
; APPLICANT: ROSSIO, PATRICIA
; TITLE OF INVENTION: THE ULTIMATE TREATMENT OF PSORIASIS
; FILE REFERENCE: 034227-098
; CURRENT APPLICATION NUMBER: US/11/249,305
; CURRENT FILING DATE: 2005-10-14
; PRIOR APPLICATION NUMBER: PCT/FR04/000934
; PRIOR FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: FR 0304773
; PRIOR FILING DATE: 2003-04-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 3
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-249-305-3

Query Match 75.3%; Score 12.8; DB 7; Length 500;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCCGA 16
Db 77 GCCCAGCGCTTGGCCGA 62

RESULT 21
US-11-249-305-4
; Sequence 4, Application US/11249305
; Publication No. US20060115836A1
```

```
; GENERAL INFORMATION:
; APPLICANT: TRINQUET, VALERIE
; APPLICANT: ROSSIO, PATRICIA
; APPLICANT: FOGEL, PAUL
; TITLE OF INVENTION: UTILIZATION OF FANCC, DAD1, GRIM19 AND HADH1 GENES FOR
; FILE REFERENCE: 034227-098
; CURRENT APPLICATION NUMBER: US/11/249,305
; CURRENT FILING DATE: 2005-10-14
; PRIOR APPLICATION NUMBER: PCT/FR04/000934
; PRIOR FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: FR 0304773
; PRIOR FILING DATE: 2003-04-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 4
; LENGTH: 500
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-249-305-4

Query Match 75.3%; Score 12.8; DB 7; Length 500;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCCGA 16
Db 424 GCCCAGCGCTTGGCCGA 439

RESULT 22
US-11-217-529-82759
; Sequence 82759, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82759
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82759

Query Match 75.3%; Score 12.8; DB 7; Length 693;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGGCCGAG 17
Db 534 CCCACGCTTGGCCGAG 549

RESULT 23
US-10-953-349-20941/c
; Sequence 20941, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
```

; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20941
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20941

Query Match 75.3%; Score 12.8; DB 6; Length 726;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17
||| ||||| ||||| |||||
Db 395 CCCGCGCTTGCGCGAG 380

RESULT 24
US-11-217-529-82116/c
; Sequence 82116, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82116
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82116

Query Match 75.3%; Score 12.8; DB 7; Length 777;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17
||| ||||| ||||| |||||
Db 125 CCCACACTTGCCCAAG 110

RESULT 25
US-11-217-529-1630
; Sequence 1630, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1630

; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1630

Query Match 75.3%; Score 12.8; DB 7; Length 1086;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCAGCGTTGGCCGA 16
||| ||||| ||||| |||||
Db 24 GCCCAGCGTTGGCGAGA 39

RESULT 26
US-10-953-349-25892/c
; Sequence 25892, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25892
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-25892

Query Match 75.3%; Score 12.8; DB 6; Length 1093;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17
||| ||||| ||||| |||||
Db 805 CCCACGCTTGCCCGAG 790

RESULT 27
US-10-953-349-33000/c
; Sequence 33000, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33000
; LENGTH: 1124
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33000

Query Match 75.3%; Score 12.8; DB 6; Length 1124;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17
||| ||||| ||||| |||||
Db 58 CCCATGCGTGCCCGAG 43

RESULT 28
US-10-953-349-35987
; Sequence 35987, Application US/10953349

```
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35987
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-35987

Query Match          75.3%; Score 12.8; DB 6; Length 1541;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GCCCAGCGTTGGCCGA 16
        ||||| ||||| |||||
Db      1372 GCCCAATCTGGCCGA 1387

RESULT 29
US-10-953-349-33688/c
; Sequence 33688, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33688
; LENGTH: 1824
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33688

Query Match          75.3%; Score 12.8; DB 6; Length 1824;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CCCACGCTTGCCCGAG 17
        ||||| ||||| |||||
Db      1088 CCCATGCTAGGCCGAG 1073

RESULT 30
US-11-293-697-1055
; Sequence 1055, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1055
; LENGTH: 2525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1055
```

```
Query Match          75.3%; Score 12.8; DB 7; Length 2525;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CCCACGCTTGCCCGAG 17
        ||||| ||||| |||||
Db      110 CCCACGCTTGCCCGAG 125

RESULT 31
US-11-293-697-1132/c
; Sequence 1132, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1132
; LENGTH: 2557
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1132

Query Match          75.3%; Score 12.8; DB 7; Length 2557;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CCCACGCTTGCCCGAG 17
        ||||| ||||| |||||
Db      876 CCCACGCTTGCCCGAG 861

RESULT 32
US-11-293-697-845
; Sequence 845, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 845
; LENGTH: 2593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-845

Query Match          75.3%; Score 12.8; DB 7; Length 2593;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 CCCACGCTTGCCCGAG 17
        ||||| ||||| |||||
Db      1291 CCCTGCTTGCCCGAG 1306

RESULT 33
US-11-293-697-1461/c
; Sequence 1461, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
```

```
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1461
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1461

Query Match          75.3%; Score 12.8; DB 7; Length 2641;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CCCACGCTTGCCCGAG 17
    ||||| ||||| |||||
Db  118 CCCAAGCTTGGCTGAG 133

RESULT 36
US-11-293-697-291
; Sequence 291, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 291
; LENGTH: 3095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-291

Query Match          75.3%; Score 12.8; DB 7; Length 3095;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CCCACGCTTGCCCGAG 17
    ||||| ||||| |||||
Db  2952 CCCACCTTGGCCGAG 2967

RESULT 37
US-11-254-185-1
; Sequence 1, Application US/11254185
; Publication No. US20060099625A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; APPLICANT: SCHUMAN, MARC
; TITLE OF INVENTION: MT-SPI SERINE PROTEASE
; FILE REFERENCE: 28644-701.302
; CURRENT APPLICATION NUMBER: US/11/254,185
; CURRENT FILING DATE: 2005-10-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 3142
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)..(2601)
US-11-254-185-1

Query Match          75.3%; Score 12.8; DB 7; Length 3142;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CCCACGCTTGCCCGAG 17
    ||||| ||||| |||||
Db  2876 CCCAAGCTTGGCCGAG 2891

RESULT 38
US-11-253-869-1
; Sequence 1, Application US/11253869
; Publication No. US20060104979A1
```

```
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1497
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1497/c

Query Match          75.3%; Score 12.8; DB 7; Length 2641;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CCCACGCTTGCCCGAG 17
    ||||| ||||| |||||
Db  1006 CCCACACTTGGCAGAG 991

RESULT 34
US-11-293-697-1497/c
; Sequence 1497, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1497
; LENGTH: 2825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1497

Query Match          75.3%; Score 12.8; DB 7; Length 2825;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  1 GCCCAGCTTGCCGCA 16
    ||||| ||||| |||||
Db  1938 GCCCAGCTTGGCAGA 1923

RESULT 35
US-11-293-697-885
; Sequence 885, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 885
; LENGTH: 3039
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-885

Query Match          75.3%; Score 12.8; DB 7; Length 3039;
```

```

; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; APPLICANT: SCHUMAN, MARC
; TITLE OF INVENTION: MT-SPI POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 28644-701.303
; CURRENT APPLICATION NUMBER: US/11/253,869
; CURRENT FILING DATE: 2005-10-18
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 3142
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)..(2601)
US-11-253-869-1

Query Match      75.3%; Score 12.8; DB 7; Length 3142;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCACGCTTGGCCGAG 17
Db 2876 CCCACGCTTGGCCGAG 2891
||||| ||| |||||

RESULT 39
US-11-293-697-510
; Sequence 510, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 510
; LENGTH: 3693
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-510

Query Match      75.3%; Score 12.8; DB 7; Length 3693;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCACGCTTGGCCGAG 17
Db 1760 CCCACGCTTGGCCGAG 1775
||||| ||| |||||

RESULT 40
US-11-121-154-158
; Sequence 158, Application US/1121154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
```

```

; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 158
; LENGTH: 5202
; TYPE: DNA
; ORGANISM: Microbulbifer degradans
US-11-121-154-158

Query Match      75.3%; Score 12.8; DB 7; Length 5202;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGA 16
Db 4889 GCCCAGCTTGGCCGA 4904
||||| ||| |||||

RESULT 41
US-11-217-529-1662
; Sequence 1662, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1662
; LENGTH: 9426
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1662

Query Match      75.3%; Score 12.8; DB 7; Length 9426;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCCGA 16
Db 8285 GCCCAGCTTGGCTGA 8300
||||| ||| |||||

RESULT 42
US-10-505-928-568/c
; Sequence 568, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 568
; LENGTH: 10905
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-568

Query Match      75.3%; Score 12.8; DB 6; Length 10905;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
```


Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17
|||||
Db 774 CCCACGCTTGCCCGAG 759

RESULT 43

US-10-857-260-34
; Sequence 34, Application US/10857260
; Publication No. US20060110742A1
; GENERAL INFORMATION:
; APPLICANT: Lyons, Leslie A.
; APPLICANT: Grahn, Robert
; APPLICANT: Erdman, Carolyn
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Carrier Tests for Polycystic Kidney Disease in the Cat
; FILE REFERENCE: 023070-146800US
; CURRENT APPLICATION NUMBER: US/10/857,260
; CURRENT FILING DATE: 2004-05-28
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 27419
; TYPE: DNA
; ORGANISM: Felis catus
; FEATURE:
; OTHER INFORMATION: GenBank Accession No. AC145332.28 genomic DNA for
; OTHER INFORMATION: PKD1 27419 bp contig
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(27419)
; OTHER INFORMATION: n = g, a, c or t

US-10-857-260-34

Query Match 75.3%; Score 12.8; DB 6; Length 27419;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 CCCACGCTTGCCCGAG 17
|||||
Db 16104 CCCCCTTGCCCGAG 16119

RESULT 44

US-10-505-928-151
; Sequence 151, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 151
; LENGTH: 128361
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-151

Query Match 75.3%; Score 12.8; DB 6; Length 128361;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCCACGCTTGCCCGAG 17
|||||
Db 46066 CCCACACTTGCCCAAG 46081

RESULT 45

US-10-489-730-10
; GENERAL INFORMATION:
; APPLICANT: Melino, Gennaro
; APPLICANT: Hayes, Ian
; APPLICANT: de Laurenzi, Vincenzo
; APPLICANT: Barcarolli, Daniela
; APPLICANT: Candi, Eleonora
; APPLICANT: Bernassola, Francesca
; APPLICANT: Tobler, Andreas
; APPLICANT: Novak, Urban
; TITLE OF INVENTION: Human Delta-N p73 Molecules and Uses Thereof
; FILE REFERENCE: 19319.002
; CURRENT APPLICATION NUMBER: US/10/489,730
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: PCT/GB02/04238
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/322,436
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 39
; OTHER INFORMATION: reverse complement of exons 14 through 1 as reported in GenBank A
US-10-489-730-10

Query Match 75.3%; Score 12.8; DB 6; Length 138941;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCCCGA 16

|||||
Db 116980 GCCCACGCGGCGCA 116995

RESULT 46

US-11-169-140-82
; Sequence 82, Application US/11169140
; Publication No. US20060099150A1
; GENERAL INFORMATION:
; APPLICANT: ARIZEKE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven
; TITLE OF INVENTION: METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE AGENTS ACROSS CE
; FILE REFERENCE: 11474-037-999
; CURRENT APPLICATION NUMBER: US/11/169,140
; CURRENT FILING DATE: 2005-06-27
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/267,601
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/248,478
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/237,929
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 09/949,039
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 09/969,748
; PRIOR FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 82
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Murinae gen. sp.
US-11-169-140-82

Query Match 72.9%; Score 12.4; DB 7; Length 39;
Best Local Similarity 92.9%; Pred. No. 1.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCCACGCTTGCC 14

Db 1 GCCCAAGCTTGCC 14
||||| |||||||

RESULT 47

US-11-169-140-87
; Sequence 87, Application US/11169140
; Publication No. US20060099150A1
; GENERAL INFORMATION:
; APPLICANT: ARIZEKE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven

; TITLE OF INVENTION: METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE AGENTS ACROSS CELL BARRIERS
; FILE REFERENCE: 11474-037-999
; CURRENT FILING DATE: 2005-06-27
; PRIOR FILING DATE: 2005-06-27
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/267,601
; PRIOR APPLICATION NUMBER: US 60/248,819
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/248,478
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/237,929
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 09/949,039
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 09/969,748
; PRIOR FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Murinae gen. sp.

US-11-169-140-87

Query Match 72.9%; Score 12.4; DB 7; Length 39;
Best Local Similarity 92.9%; Pred. No. 1.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCC 14
||||| |||||||

Db 1 GCCCAAGCTTGCC 14

RESULT 48

US-11-169-140-95
; Sequence 95, Application US/11169140
; Publication No. US20060099150A1
; GENERAL INFORMATION:
; APPLICANT: ARIZEKE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven

; TITLE OF INVENTION: METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE AGENTS ACROSS CELL BARRIERS
; FILE REFERENCE: 11474-037-999
; CURRENT APPLICATION NUMBER: US/11/169,140
; CURRENT FILING DATE: 2005-06-27
; PRIOR FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 60/267,601
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/248,819
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/248,478
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/237,929
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: 09/949,039
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 09/969,748
; PRIOR FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 95
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Rattus sp.

US-11-169-140-95

Query Match 72.9%; Score 12.4; DB 7; Length 39;
Best Local Similarity 92.9%; Pred. No. 1.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGCC 14
||||| |||||||

Db 1 GCCCAAGCTTGCC 14

RESULT 49

US-10-953-349-26326
; Sequence 26326, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26326
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (477)..(477)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-26326

Query Match 72.9%; Score 12.4; DB 6; Length 506;
Best Local Similarity 92.9%; Pred. No. 2.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCACGCTTGCCG 15
||||| |||||||

Db 204 CCCATGCTTGCCG 217

RESULT 50

US-10-953-349-37335/c
; Sequence 37335, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 37335
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-37335

Query Match 72.9%; Score 12.4; DB 6; Length 1035;

Best Local Similarity 92.9%; Pred. No. 2.2e+02; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 1;

QY 1 GCCCAGCTTGGCC 14
|||||
Db 324 GCCCAGCTTGGCC 311

RESULT 51

US-11-217-529-272/c
; Sequence 272, Application US/11217529
; Publication No. US2006009612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHO
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 272
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-272

Query Match 72.9%; Score 12.4; DB 7; Length 1242;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCC 14
|||||
Db 120 GCCCAGCTTGGCC 107

RESULT 52

US-10-471-571A-2011
; Sequence 2011, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P02692WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 2011
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2011

Query Match 72.9%; Score 12.4; DB 6; Length 1305;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGCTTGGCC 14
|||||
Db 75 GCCCAGCTTGGCC 88

RESULT 53

US-10-953-349-35503/c

; Sequence 35503, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35503
; LENGTH: 1423
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (513)..(513)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-35503

Query Match 72.9%; Score 12.4; DB 6; Length 1423;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCACGCTTGGCGA 16.
|||||
Db 81 CCACGCTAGGCCGA 68

RESULT 54

US-10-953-349-35104
; Sequence 35104, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35104
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-35104

Query Match 72.9%; Score 12.4; DB 6; Length 1429;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CACGCTTGGCCGAG 17
|||||
Db 1189 CACGATGCGCGAG 1202

RESULT 55

US-11-293-697-3/c
; Sequence 3, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3

```
; LENGTH: 1685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-3

Query Match      72.9%; Score 12.4; DB 7; Length 1685;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 CACGCTTGCCGAG 17
Db      601 CACCTTGGCCGAG 588

RESULT 56
US-11-293-697-559/c
; Sequence 559, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 559
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-559

Query Match      72.9%; Score 12.4; DB 7; Length 2175;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCC 14
Db      817 GCCCAGCGTTGGCC 804

RESULT 57
US-11-293-697-322
; Sequence 322, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 322
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-322

Query Match      72.9%; Score 12.4; DB 7; Length 2299;
Best Local Similarity 92.9%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 CACGCTTGCCGAG 17
Db      417 CACGCTTGCCGAG 430

RESULT 58
US-11-293-697-1089
; Sequence 1089, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1089
; LENGTH: 2728
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1089

Query Match      72.9%; Score 12.4; DB 7; Length 2728;
Best Local Similarity 92.9%; Pred. No. 2.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCC 14
Db      573 GCCCAGCGTTGGCC 586

RESULT 59
US-11-293-697-1682/c
; Sequence 1682, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1682
; LENGTH: 2750
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1682

Query Match      72.9%; Score 12.4; DB 7; Length 2750;
Best Local Similarity 92.9%; Pred. No. 2.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GCCCAGCGTTGGCC 14
Db      187 GCCCAGCGTTGGCC 174

RESULT 60
US-10-511-937-576/c
; Sequence 576, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
```

```

; SEQ ID NO 5
; LENGTH: 16032

```

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding MSP1D5D6
US-11-259-950-11

Query Match 71.8%; Score 12.2; DB 7; Length 522;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCGGAG 17
||| ||||| |||||
Db 293 GCCAGCGCTTGGCGGCG 309

RESULT 65
US-11-217-997-9
; Sequence 9, Application US/11217997
; Publication No. US20060111561A1
; GENERAL INFORMATION:
; APPLICANT: Valerie L. Gerlach
; APPLICANT: Elma R. Fernandes
; APPLICANT: Richard A. Shimkets
; APPLICANT: Meera Patturajan
; APPLICANT: Vladimir Y. Gusev
; APPLICANT: Stacie (Casman) Navara
; APPLICANT: Velizar T. Tchernev
; APPLICANT: David W. Anderson
; APPLICANT: Xiaojia (Saaha) Guo
; APPLICANT: Luca Rastelli
; APPLICANT: Mei Zhong
; APPLICANT: Muralidhara Padigaru
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODED THEREBY
; FILE REFERENCE: Cura 551 CIP
; CURRENT APPLICATION NUMBER: US/11/217,997
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: 10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 9
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (29)..(514)
US-11-217-997-9

Query Match 71.8%; Score 12.2; DB 7; Length 522;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGGCGGAG 17
||| ||||| |||||
Db 62 GCCCGCGCTGTGCGGAG 78

RESULT 66
US-11-217-997-35
; Sequence 35, Application US/11217997
; Publication No. US20060111561A1
; GENERAL INFORMATION:
; APPLICANT: Valerie L. Gerlach
; APPLICANT: Elma R. Fernandes
; APPLICANT: Richard A. Shimkets
; APPLICANT: Meera Patturajan
; APPLICANT: Vladimir Y. Gusev
; APPLICANT: Stacie (Casman) Navara
; APPLICANT: Velizar T. Tchernev
; APPLICANT: David W. Anderson
; APPLICANT: Xiaojia (Saaha) Guo
; APPLICANT: Luca Rastelli
; APPLICANT: Mei Zhong
; APPLICANT: Muralidhara Padigaru
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODED THEREBY
; FILE REFERENCE: Cura 551 CIP
; CURRENT APPLICATION NUMBER: US/11/217,997
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: 10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 35
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(520)
US-11-217-997-35

Query Match 71.8%; Score 12.2; DB 7; Length 522;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTTGGCGGAG 17
||| ||||| |||||
Db 62 GCCCGCGCTGTGCGGAG 78

RESULT 67
US-11-330-650-2/c
; Sequence 2, Application US/11330650
; Publication No. US2006094055A1

```
; GENERAL INFORMATION:
; APPLICANT: Ding, Shou-wei
; APPLICANT: Li, Hong-wei
; APPLICANT: Li, Wan-xiang
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: RNA Silencing in Animals as an Antiviral Defense
; FILE REFERENCE: 023070-124100US
; CURRENT APPLICATION NUMBER: US/11/330,650
; CURRENT FILING DATE: 2006-01-11
; PRIOR APPLICATION NUMBER: US/10/150,283
; PRIOR FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 540
; TYPE: RNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; OTHER INFORMATION: F protein
US-11-330-650-2

Query Match          71.8%; Score 12.2; DB 7; Length 540;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTTGGCCGAG 17
   ||||| ||||| ||||| |||||
Db 324 GCCCAGCTAGGCCGAG 308

RESULT 68
US-10-488-619-1465/c
; Sequence 1465, Application US/10488619
; Publication No. US2006009578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1465
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-1465

Query Match          71.8%; Score 12.2; DB 6; Length 562;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTTGGCCGAG 17
   ||||| ||||| ||||| |||||
Db 209 GCCCAGCGCTTGGCCGG 193

RESULT 69
US-11-259-950-5
; Sequence 5, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; APPLICANT: Sligar, Stephen G.
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
; PRIOR FILING DATE: 2004-10-27
; PRIOR APPLICATION NUMBER: US 60/536,281
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US 10/465,789
```

```
; GENERAL INFORMATION:
; APPLICANT: Ding, Shou-wei
; APPLICANT: Li, Hong-wei
; APPLICANT: Li, Wan-xiang
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: RNA Silencing in Animals as an Antiviral Defense
; FILE REFERENCE: 023070-124100US
; CURRENT APPLICATION NUMBER: US/11/330,650
; CURRENT FILING DATE: 2006-01-11
; PRIOR APPLICATION NUMBER: US/10/150,283
; PRIOR FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 540
; TYPE: RNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; OTHER INFORMATION: F protein
US-11-330-650-2

Query Match          71.8%; Score 12.2; DB 7; Length 619;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTTGGCCGAG 17
   ||||| ||||| ||||| |||||
Db 390 GCCCAGCGCTTGGCCGG 406

RESULT 70
US-10-953-349-14927/c
; Sequence 14927, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14927
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-14927

Query Match          71.8%; Score 12.2; DB 6; Length 633;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGCTTGGCCGAG 17
   ||||| ||||| ||||| |||||
Db 97 GCCCAGCGCTTGGCCGAG 81

RESULT 71
US-11-259-950-3
; Sequence 3, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; APPLICANT: Sligar, Stephen G.
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
; PRIOR FILING DATE: 2004-10-27
; PRIOR APPLICATION NUMBER: US 60/536,281
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US 10/465,789
```


; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 60/252,233
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 11/033,489
; PRIOR FILING DATE: 2005-01-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding MSPI
US-11-259-950-3

Query Match 71.8%; Score 12.2; DB 7; Length 654;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| ||||| ||||| |||
Db 425 GCCCAGCGCTGGCCGCG 441

RESULT 72

US-10-488-619-2847
; Sequence 2847, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2847
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2847

Query Match 71.8%; Score 12.2; DB 6; Length 655;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| ||||| ||||| |||
Db 483 GCCCAGCGCTGGTCAG 499

RESULT 73

US-11-217-529-566
; Sequence 566, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; TITLE OF INVENTION: ASHIKARI, TOSHIOHIKO
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 566
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-566

Query Match 71.8%; Score 12.2; DB 7; Length 681;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| ||||| ||||| |||
Db 378 GCCCAGCGTTTGGCCAAG 394

RESULT 74

US-11-259-950-1
; Sequence 1, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
; PRIOR FILING DATE: 2004-10-27
; PRIOR APPLICATION NUMBER: US 60/536,281
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US 10/465,789
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 60/252,233
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 11/033,489
; PRIOR FILING DATE: 2005-01-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-259-950-1

Query Match 71.8%; Score 12.2; DB 7; Length 762;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| ||||| ||||| |||
Db 535 GCCCAGCGCTGGCCGCG 551

RESULT 75

US-10-953-349-28486
; Sequence 28486, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28486
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Triticum aestivum

US-10-953-349-28486

Query Match 71.8%; Score 12.2; DB 6; Length 798;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
| | | | | | | | | | | | | | | | | | | |
Db 687 GTCCAGCGCTGGCCGAG 703

RESULT 76

US-10-196-749-487
; Sequence 487, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 487
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Homo Sapien

US-10-196-749-487

Query Match 71.8%; Score 12.2; DB 6; Length 843;
Best Local Similarity 82.4%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
| | | | | | | | | | | | | | | | | | | |
Db 767 GGCCACCTCTTGGCAGAG 783

RESULT 77

US-10-953-349-20319
; Sequence 20319, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20319
; LENGTH: 868
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20319

Query Match 71.8%; Score 12.2; DB 6; Length 868;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
| | | | | | | | | | | | | | | | | | | |
Db 519 GCCCGCGCTTGGCCGAG 535

RESULT 78

US-11-217-997-7
; Sequence 7, Application US/11217997
; Publication No. US20060111561A1
; GENERAL INFORMATION:
; APPLICANT: Valerie L. Gerlach
; APPLICANT: Elma R. Fernandes
; APPLICANT: Richard A. Shimkets
; APPLICANT: Meera Patturajan
; APPLICANT: Vladimir Y. Gusev
; APPLICANT: Stacie (Casman) Navara
; APPLICANT: Velizar T. Tchernev
; APPLICANT: David W. Anderson
; APPLICANT: Xiaojia (Sasha) Guo
; APPLICANT: Luca Rastelli
; APPLICANT: Mei Zhong
; APPLICANT: Muralidhara Padigaru
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODED THEREBY
; FILE REFERENCE: Cura 551 CIP
; CURRENT APPLICATION NUMBER: US/11/217,997
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: 10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: CuraSequist version 0.1
; SEQ ID NO 7
; LENGTH: 877
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (25) .. (534)
US-11-217-997-7

```

Query Match	71.8%	Score 12.2;	DB 7;	Length 877;
Best Local Similarity	82.4%	Pred. No. 2.8e+02;		
Matches 14;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

```

RESULT 79
US-10-953-349-28667/c
; Sequence 28667, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28667
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-28667

```

Query Match	71.8%;	Score 12.2;	DB 6;	Length 945;
Best Local Similarity	82.4%;	Pred. No. 2.8e+02;		
Matches 14;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

```

RESULT 80
US-11-256-428-62
/ Sequence 62, Application US/11256428
/ Publication No. US20060095987A1
/ GENERAL INFORMATION:
/ APPLICANT: Niblett, Charles L.
/ TITLE OF INVENTION: Methods and Materials
/ TITLE OF INVENTION: Pathogens of Plants
/ FILE REFERENCE: VEN-100
/ CURRENT APPLICATION NUMBER: US/11/256,428
/ CURRENT FILING DATE: 2005-10-21
/ NUMBER OF SEQ ID NOS: 69
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 62
/ LENGTH: 949
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Heterodera glycines r
US-11-256-428-62

```

Query Match	71.8%;	Score 12.2;	DB 7;	Length 949;
Best Local Similarity	82.4%;	Pred. No. 2.8e+03;		
Matches 14;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps 0;

RESULT 81
US-10-953-349-27232

```

; Sequence 27232, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27232
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-27232

```

Query Match	71.8%;	Score 12.2;	DB 6;
Best Local Similarity	82.4%;	Pred. No. 2.8e+02;	Length 993;
Matches 14;	Conservative	0;	Mismatches 3;
			Indels 0;
			Gaps 0;

```

RESULT 82
US-10-953-349-3387
; Sequence 3387, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3387
; LENGTH: 1038
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3387

```

Query Match	71.8%	Score 12.2;	DB 6;	Length 1038;
Best Local Similarity	82.4%;	Pred. No. 2.8e+03;		
Matches 14:	Conservative	0;	Mismatches 3;	Indels 0;
	Conservative	0;	Mismatches 3;	Indels 0;

```

RESULT 83
US-10-953-349-35759/c
? Sequence 35759, Application US/10953349
? Publication No. US20060107345A1
? GENERAL INFORMATION:
? APPLICANT: ALEXANDROV, Nikolai et al.
? TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
? TITLE OF INVENTION: ENCODED THERBY
? FILE REFERENCE: 2750-1579PUS2
? CURRENT APPLICATION NUMBER: US/10/953,349
? CURRENT FILING DATE: 2004-09-30
? NUMBER OF SEQ ID NOS: 40252
? SOFTWARE: PatentIn version 3.3
? SEQ ID NO 35759
? LENGTH: 1233
? TYPE: DNA
? ORGANISM: Zea mays subsp. mays
US-10-953-349-35759

```

```
Query Match          71.8%; Score 12.2; DB 6; Length 1233;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
Db 774 GCCCCCGGAGGCCGAG 758

RESULT 84
US-11-254-252-1/c
; Sequence 1, Application US/11254252
; Publication No. US20060110755A1
; GENERAL INFORMATION:
; APPLICANT: Duke, Richard C.
; APPLICANT: Franzusoff, Alex
; APPLICANT: Haller, Aurelia
; APPLICANT: King, Thomas H.
; TITLE OF INVENTION: YEAST-BASED THERAPEUTIC FOR CHRONIC HEPATITIS C INFECTION
; FILE REFERENCE: 3923-12
; CURRENT APPLICATION NUMBER: US/11/254,252
; CURRENT FILING DATE: 2005-10-18
; PRIOR APPLICATION NUMBER: US 60/620,158
; PRIOR FILING DATE: 2004-10-18
; PRIOR APPLICATION NUMBER: US 10/738,646
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 60/434,163
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: recombinant fusion protein construct
US-11-254-252-1

Query Match          71.8%; Score 12.2; DB 7; Length 1233;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
Db 1128 GCCCAGCTAGGCCGAG 1112

RESULT 85
US-11-259-950-7
; Sequence 7, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; APPLICANT: Sligar, Stephen G.
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
; PRIOR FILING DATE: 2004-10-27
; PRIOR APPLICATION NUMBER: US 60/536,281
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US 10/465,789
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 60/252,233
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 11/033,489
; PRIOR FILING DATE: 2005-01-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding MSP2TEV

Query Match          71.8%; Score 12.2; DB 7; Length 1275;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
Db 452 GCCAGCGCTTGCCGCG 468

RESULT 87
US-11-259-950-9
; Sequence 9, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; APPLICANT: Sligar, Stephen G.
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
```

```
; SEQ ID NO 7
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding MSP2 with short linker
US-11-259-950-7

Query Match          71.8%; Score 12.2; DB 7; Length 1260;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
Db 425 GCCAGCGCTTGCCGCG 441

RESULT 86
US-11-259-950-77
; Sequence 77, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; APPLICANT: Sligar, Stephen G.
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
; PRIOR FILING DATE: 2004-10-27
; PRIOR APPLICATION NUMBER: US 60/536,281
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US 10/465,789
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 60/252,233
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 11/033,489
; PRIOR FILING DATE: 2005-01-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding MSP2TEV

US-11-259-950-77

Query Match          71.8%; Score 12.2; DB 7; Length 1275;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCAGCGTTGGCCGAG 17
Db 452 GCCAGCGCTTGCCGCG 468

RESULT 87
US-11-259-950-9
; Sequence 9, Application US/11259950
; Publication No. US20060088524A1
; GENERAL INFORMATION:
; APPLICANT: Morrissey, James
; APPLICANT: Pureza, Vincent
; APPLICANT: Sligar, Stephen G.
; TITLE OF INVENTION: Tissue Factor Compositions and Methods
; FILE REFERENCE: 46-04
; CURRENT APPLICATION NUMBER: US/11/259,950
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US 60/622,737
```

; PRIOR FILING DATE: 2004-10-27
; PRIOR APPLICATION NUMBER: US 60/536,281
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US 10/465,789
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 60/252,233
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 11/033,489
; PRIOR FILING DATE: 2005-01-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding MSP2L (with long linker)
US-11-259-950-9

Query Match 71.8%; Score 12.2; DB 7; Length 1282;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| |||||
Db 423 GCCAGCGCTGGCCGCG 439

RESULT 88
US-10-953-349-35397/c
; Sequence 35397, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35397
; LENGTH: 1295
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-35397

Query Match 71.8%; Score 12.2; DB 6; Length 1295;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| |||||
Db 247 GCCCGCGCTGGCCGAG 231

RESULT 89
US-10-953-349-27665/c
; Sequence 27665, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27665
; LENGTH: 1378

; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-27665

Query Match 71.8%; Score 12.2; DB 6; Length 1378;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| |||||
Db 940 GCCCGCGCTGGCCGCG 924

RESULT 90
US-10-953-349-28083
; Sequence 28083, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28083
; LENGTH: 1409
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-28083

Query Match 71.8%; Score 12.2; DB 6; Length 1409;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| |||||
Db 1340 GCCCATGATTGGCCGTG 1356

RESULT 91
US-10-953-349-35361
; Sequence 35361, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35361
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-35361

Query Match 71.8%; Score 12.2; DB 6; Length 1605;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCCAGCGCTGGCCGAG 17
||| |||||
Db 234 GCCCGCGCTGGCCGTG 250

RESULT 92
US-10-505-928-653
; Sequence 653, Application US/10505928
; Publication No. US20060089532A1

```
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 653
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-653

Query Match      71.8%; Score 12.2; DB 6; Length 1645;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GCCACGCTTGCCGAG 17
      ||| ||||| |||
Db      1569  GCCTACGCTTGCCAG 1585

RESULT 93
US-11-293-697-981
; Sequence 981, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 981
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-981

Query Match      71.8%; Score 12.2; DB 7; Length 1736;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GCCACGCTTGCCGAG 17
      ||| ||||| |||
Db      620  GTCACGCTGGGCTAG 636

RESULT 94
US-10-953-349-34645/c
; Sequence 34645, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34645
; LENGTH: 1859
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34645

Query Match      71.8%; Score 12.2; DB 6; Length 1859;
Best Local Similarity 82.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GCCACGCTTGCCGAG 17
      ||| ||||| |||
Db      1375  GCTAGCGCTTGCCGAG 1359

RESULT 95
US-11-254-252-13/c
; Sequence 13, Application US/11254252
; Publication No. US20060110755A1
; GENERAL INFORMATION:
; APPLICANT: Duke, Richard C.
; APPLICANT: Franzusoff, Alex
; APPLICANT: Haller, Aurelia
; APPLICANT: King, Thomas H
; TITLE OF INVENTION: YEAST-BASED THERAPEUTIC FOR CHRONIC HEPATITIS C INFECTION
; FILE REFERENCE: 3923-12
; CURRENT APPLICATION NUMBER: US/11/254,252
; CURRENT FILING DATE: 2005-10-18
; PRIOR APPLICATION NUMBER: US 60/620,158
; PRIOR FILING DATE: 2004-10-18
; PRIOR APPLICATION NUMBER: US 10/738,646
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 60/434,163
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: recombinant fusion protein construct
US-11-254-252-13

Query Match      71.8%; Score 12.2; DB 7; Length 1908;
Best Local Similarity 82.4%; Pred. No. 2.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GCCACGCTTGCCGAG 17
      ||| ||||| |||
Db      339  GCCCCAGCTAGGCCGAG 323

RESULT 96
US-11-217-529-79927/c
; Sequence 79927, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79927
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79927

Query Match      71.8%; Score 12.2; DB 7; Length 1941;
```

Best Local Similarity 82.4%; Pred. NO. 2.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGCCGAG 17
Db 408 GCCCAGCGCTTGACGAG 392

RESULT 97

US-11-293-697-2383
; Sequence 2383, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2383
; LENGTH: 2128
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-2383

Query Match 71.8%; Score 12.2; DB 7; Length 2128;
Best Local Similarity 82.4%; Pred. NO. 2.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGCCGAG 17
Db 800 GCCCAGCGCTTGCCGAG 816

RESULT 98

US-11-293-697-1876/C
; Sequence 1876, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1876
; LENGTH: 2182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1876

Query Match 71.8%; Score 12.2; DB 7; Length 2182;
Best Local Similarity 82.4%; Pred. NO. 2.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGCCGAG 17
Db 1523 GCCCAGCGCTTGCCCTAG 1507

RESULT 99

US-11-293-697-40
; Sequence 40, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA

FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-40

Query Match 71.8%; Score 12.2; DB 7; Length 2208;
Best Local Similarity 82.4%; Pred. NO. 2.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGCCGAG 17
Db 659 GCCCAGCGCTTGCTGAG 675

RESULT 100

US-11-293-697-40/C
; Sequence 40, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-40

Query Match 71.8%; Score 12.2; DB 7; Length 2208;
Best Local Similarity 82.4%; Pred. NO. 2.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCCCAGCGCTTGCCGAG 17
Db 1770 GCCCAGCGCATGCCGGG 1754

Search completed: June 10, 2006, 15:37:52
Job time : 14.2167 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:18:44 ; Search time 518.067 Seconds
(without alignments)
2345.260 Million cell updates/sec

Title: US-10-600-816-32

Perfect score: 19

Sequence: 1 cttacttcattagctttg 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	100.0	603	2	BD233463
C 2	19	100.0	1212	2	BD2029699
C 3	19	100.0	1212	2	AR341505
C 4	19	100.0	1619	2	AX078375
C 5	19	100.0	1718	2	BD233473
C 6	19	100.0	1826	5	AK172760
C 7	19	100.0	2290	2	C0723177
C 8	19	100.0	2295	5	BC003665
C 9	19	100.0	2297	5	AF506289
C 10	19	100.0	2302	5	AX930411
C 11	19	100.0	2302	5	AF095448
C 12	19	100.0	2446	2	BD156680
C 13	19	100.0	2446	2	AX877483
C 14	19	100.0	2446	5	AX001761
C 15	19	100.0	2456	2	C0981495
C 16	19	100.0	2456	2	DD210040
C 17	19	100.0	2456	2	AX549168
C 18	19	100.0	3057	5	AK122672

AX188348 Sequence
C0894732 Sequence
AC007688 Homo sapi
AL954230 Pan trogl
AF064865 Homo sapi
AC087481 Homo sapi
AL954231 Pan trogl
AL954232 Homo sapi
AC093102 Rattus no
AC093102 Rattus no
AL163281 Homo sapi
DQ271038 Mustela v
Z72916 S. cerevisia
Z72915 S. cerevisia
AY800241 Homo sapi
AC157840 Aotus nan
Continuation (68 o
Continuation (150
Continuation (25 o
AP001818 Homo sapi
AC150029 Sorex ara
AC167325 Loxodonta
AL121825 Human DNA
AC011354 Homo sapi
AC011334 Homo sapi
AP006464 Oryza sat
AC091155 Homo sapi
AC074120 Homo sapi
AP003856 Oryza sat
AP003556 Homo sapi
AC173916 Atelerix
AP003021 Oryza sat
AC015570 Homo sapi
AC021588 Homo sapi
AC124725 Mus muscu
AP003174 Homo sapi
AC127007 Rattus no
AC164008 Mus muscu
AC148231 Didelphis
AL845354 Mouse DNA
AL512635 Human DNA
AC011949 Homo sapi
AC022743 Homo sapi
AC097585 Sus scrof
AC146837 Canis fam
AC011947 Homo sapi
AC069476 Homo sapi
AC132501 Rattus no
AC161349 Mus muscu
AC144518 Pan trogl
AC151368 Aotus nan
AC097629 Sus scrof
AC124922 Mus muscu
AC102703 Mus muscu
AC151118 Mus muscu
AC146886 Callithri
AC113245 Mus muscu
AC146883 Callithri
AC147845 Saimiri b
AC092434 Homo sapi
AC112307 Rattus no
AC096378 Rattus no
AC095186 Rattus no
AC137204 Rattus no
AC108958 Rattus no
AC116211 Rattus no
AB219542 Campyloba
AC172882 Brassaica
Continuation (2 of
Continuation (2 of

92	17	89.5	110000	15	AE010299_40	Continuation (41 o	165	16.4	86.3	130213	4	AC144725	AC144725 Medicago
93	17	89.5	110038	4	AC155344	AC155344 Brassica	166	16.4	86.3	131582	6	AL953900	AL953900 Mouse DNA
C 94	17	89.5	122613	5	AL359642	AL359642 Human DNA	167	16.4	86.3	133908	12	AC182356	AC182356 Loxodonta
95	17	89.5	145877	6	AC155845	AC155845 Mus muscu	168	16.4	86.3	136436	12	AC152155	AC152155 Dasypoda
C 96	17	89.5	182422	12	CR450699	CR450699 Danio rer	C 169	16.4	86.3	137022	5	AC004231	AC004231 Homo sapi
C 97	17	89.5	187401	5	AC148203	AC148203 Callicebu	C 170	16.4	86.3	142372	5	AC132152	AC132152 Homo sapi
98	17	89.5	195080	12	CR354430	CR354430 Danio rer	C 171	16.4	86.3	143467	12	AC129996	AC129996 Rattus no
99	17	89.5	231443	12	AC111133	AC111133 Mus muscu	172	16.4	86.3	143579	12	AC157845	AC157845 Ornithorh
100	17	89.5	234055	6	AC158667	AC158667 Mus muscu	173	16.4	86.3	145001	11	CR381560	CR381560 Zebrafish
101	16.4	86.3	168	2	CS224837	CS224837 Sequence	174	16.4	86.3	145023	12	AC074380	AC074380 Homo sapi
C 102	16.4	86.3	352	4	AK105041	AK105041 Oryza sat	C 175	16.4	86.3	145156	5	AC010414	AC010414 Homo sapi
C 103	16.4	86.3	354	2	BD244398	BD244398 Determina	176	16.4	86.3	146746	5	AL353718	AL353718 Human DNA
C 104	16.4	86.3	407	2	AX916898	AX916898 Sequence	C 177	16.4	86.3	147105	5	AL136452	AL136452 Human DNA
C 105	16.4	86.3	407	2	BD052431	BD052431 Sequence	178	16.4	86.3	147434	5	AL1390789	AL1390789 Human DNA
C 106	16.4	86.3	407	2	AR752692	AR752692 Sequence	C 179	16.4	86.3	147695	5	AC136976	AC136976 Mus muscu
C 107	16.4	86.3	587	4	AK220991	AK220991 Arabidops	C 180	16.4	86.3	148198	5	HS438G17	HS438G17 Human DNA
C 108	16.4	86.3	623	7	BV259302	BV259302 S235P6104	181	16.4	86.3	148959	11	AC147203	AC147203 Xenopus t
C 109	16.4	86.3	634	7	BV032850	BV032850 S212P6825	182	16.4	86.3	150561	12	AC016457	AC016457 Homo sapi
C 110	16.4	86.3	659	7	BV075320	BV075320 S212P6037	183	16.4	86.3	150913	5	AC083875	AC083875 Homo sapi
C 111	16.4	86.3	698	7	BV633934	BV633934 S216P6112	184	16.4	86.3	152100	12	AC155565	AC155565 Zea mays
C 112	16.4	86.3	832	4	AK062234	AK062234 Oryza sat	C 185	16.4	86.3	152515	6	AC107744	AC107744 Mus muscu
C 113	16.4	86.3	835	4	AK098887	AK098887 Oryza sat	C 186	16.4	86.3	153833	12	AP001996	AP001996 Homo sapi
C 114	16.4	86.3	889	4	AK119723	AK119723 Oryza sat	C 187	16.4	86.3	154738	5	AC023355	AC023355 Homo sapi
C 115	16.4	86.3	1006	11	AY791388	AY791388 Characidi	C 188	16.4	86.3	156300	12	AC087739	AC087739 Homo sapi
C 116	16.4	86.3	1625	11	DQ062815	DQ062815 Hippoglos	189	16.4	86.3	156766	5	AC079597	AC079597 Homo sapi
C 117	16.4	86.3	1882	4	AY133832	AY133832 Arabidops	190	16.4	86.3	156817	5	AC009406	AC009406 Homo sapi
C 118	16.4	86.3	2022	11	DQ062814	DQ062814 Hippoglos	191	16.4	86.3	157097	12	AC027784	AC027784 Homo sapi
C 119	16.4	86.3	2082	4	AY074335	AY074335 Arabidops	C 192	16.4	86.3	157318	6	AC121099	AC121099 Mus muscu
C 120	16.4	86.3	2864	13	AK113019	AK113019 Clona int	193	16.4	86.3	157548	12	AC026637	AC026637 Homo sapi
C 121	16.4	86.3	3054	4	AK099656	AK099656 Oryza sat	194	16.4	86.3	157967	12	CR626901	CR626901 Danio rer
C 122	16.4	86.3	8285	15	MMU47023	U47023 Methanococc	195	16.4	86.3	158349	5	AC024908	AC024908 Homo sapi
C 123	16.4	86.3	12689	8	AF134199	AF134199 Expressio	196	16.4	86.3	158544	5	AC023658	AC023658 Homo sapi
C 124	16.4	86.3	20219	5	AC139015	AC139015 Homo sapi	C 197	16.4	86.3	158779	12	AL713851	AL713851 Human DNA
C 125	16.4	86.3	27825	12	AC015763	AC015763 Homo sapi	198	16.4	86.3	159667	5	AC078953	AC078953 Homo sapi
C 126	16.4	86.3	34217	5	HS126G10	Z82184 Human DNA s	C 199	16.4	86.3	160191	6	AC154240	AC154240 Mus muscu
C 127	16.4	86.3	34537	5	AL161428	AL161428 Human DNA	200	16.4	86.3	162401	12	CR385091	CR385091 Danio rer
C 128	16.4	86.3	37649	5	AP004722	AP004722 Homo sapi	201	16.4	86.3	163006	4	AP003284	AP003284 Oryza sat
C 129	16.4	86.3	39456	15	AE000667	AE000667 Aquifex a	C 202	16.4	86.3	164837	5	AC079603	AC079603 Homo sapi
C 130	16.4	86.3	42210	5	AC112724	AC112724 Homo sapi	203	16.4	86.3	166192	5	AC096737	AC096737 Homo sapi
C 131	16.4	86.3	42983	6	BX293999	BX293999 Mouse DNA	204	16.4	86.3	166830	12	AC173280	AC173280 Bos tauru
C 132	16.4	86.3	43087	5	AP005155	AP005155 Homo sapi	C 205	16.4	86.3	168526	6	AC122014	AC122014 Mus muscu
C 133	16.4	86.3	52675	12	BX470179_3	Continuation (4 of	C 206	16.4	86.3	169512	5	AC010996	AC010996 Homo sapi
C 134	16.4	86.3	53394	5	AC013276	AC013276 Homo sapi	207	16.4	86.3	169819	12	AC074368	AC074368 Homo sapi
C 135	16.4	86.3	73878	12	AC011600_3	Continuation (4 of	C 208	16.4	86.3	171554	6	AC142228	AC142228 Mus muscu
C 136	16.4	86.3	76133	5	HSU71148	U71148 Human Xq28	C 209	16.4	86.3	171720	6	AC131775	AC131775 Mus muscu
C 137	16.4	86.3	77452	4	AP004528	AP004528 Lotus cor	C 210	16.4	86.3	171937	5	AC008442	AC008442 Homo sapi
C 138	16.4	86.3	80828	11	AL672211	AL672211 Zebrafish	C 211	16.4	86.3	172004	6	AC173113	AC173113 Mus muscu
C 139	16.4	86.3	82710	12	AC140105	AC140105 Medicago	C 212	16.4	86.3	172300	5	AC010254	AC010254 Homo sapi
C 140	16.4	86.3	90472	4	ART6114	AL391710 Arabidops	C 213	16.4	86.3	172637	5	AC012572	AC012572 Homo sapi
C 141	16.4	86.3	95660	12	AC131181	AC131181 Homo sapi	214	16.4	86.3	172655	11	BX927400	BX927400 Zebrafish
C 142	16.4	86.3	96753	12	AC130947_4	Continuation (5 of	215	16.4	86.3	172658	5	AC093580	AC093580 Homo sapi
C 143	16.4	86.3	103421	5	AC004920	AC004920 Homo sapi	C 216	16.4	86.3	172770	12	AC148425	AC148425 Ornithorh
C 144	16.4	86.3	110000	4	AP007155_23	Continuation (24 o	C 217	16.4	86.3	173394	5	AC025040	AC025040 Homo sapi
C 145	16.4	86.3	110000	4	AP008207_270	Continuation (271 o	218	16.4	86.3	173742	12	AC068626	AC068626 Homo sapi
C 146	16.4	86.3	110000	12	AC011600_2	Continuation (3 of	C 219	16.4	86.3	174860	5	CNS01R17	AL163760 Human chr
C 147	16.4	86.3	110000	12	AC111332_1	Continuation (2 of	C 220	16.4	86.3	175216	11	BX276115	BX276115 Zebrafish
C 148	16.4	86.3	110000	12	AL359083_0	AL359083 Homo sapi	C 221	16.4	86.3	175361	6	AC122500	AC122500 Mus muscu
C 149	16.4	86.3	110000	12	BX470179_0	BX470179 Homo sapi	222	16.4	86.3	175466	5	AL607077	AL607077 Human DNA
C 150	16.4	86.3	110000	15	CP000095_13	Continuation (14 o	C 223	16.4	86.3	175583	12	AC083850	AC083850 Homo sapi
C 151	16.4	86.3	110000	15	CP000099_21	Continuation (22 o	224	16.4	86.3	175599	12	AC138670	AC138670 Mus muscu
C 152	16.4	86.3	110000	15	AE010299_02	Continuation (3 of	225	16.4	86.3	175712	6	AC102756	AC102756 Mus muscu
C 153	16.4	86.3	110000	15	AE010299_03	Continuation (4 of	226	16.4	86.3	177356	12	AC115206	AC115206 Rattus no
C 154	16.4	86.3	110000	15	BA000016_18	Continuation (19 o	227	16.4	86.3	177414	12	AC097154	AC097154 Rattus no
C 155	16.4	86.3	110000	15	BX950851_33	Continuation (34 o	C 228	16.4	86.3	177910	12	AC018357	AC018357 Homo sapi
C 156	16.4	86.3	111104	5	AC117389	AC117389 Homo sapi	229	16.4	86.3	177946	12	AC146769	AC146769 Carollia
C 157	16.4	86.3	111507	5	HS238G17	Z98753 Human DNA s	C 230	16.4	86.3	178081	6	AC132330	AC132330 Mus muscu
C 158	16.4	86.3	111560	5	AC016940	AC016940 Homo sapi	C 231	16.4	86.3	178328	12	AC019265	AC019265 Homo sapi
C 159	16.4	86.3	111909	12	AC162145	AC162145 Loxodonta	232	16.4	86.3	178510	11	AC165521	AC165521 Bos tauru
C 160	16.4	86.3	114608	12	AP007390	AP007390 Lotus cor	233	16.4	86.3	179286	11	BX897685	BX897685 Zebrafish
C 161	16.4	86.3	120709	5	AL445123	AL445123 Human DNA	C 234	16.4	86.3	179375	12	AC092971	AC092971 Homo sapi
C 162	16.4	86.3	122505	5	AC010623	AC010623 Homo sapi	235	16.4	86.3	179611	12	AC084348	AC084348 Homo sapi
C 163	16.4	86.3	122557	4	CT009492	CT009492 M.truncat	236	16.4	86.3	179767	5	AC069223	AC069223 Homo sapi
C 164	16.4	86.3	126665	12	AC115515	AC115515 Rattus no	237	16.4	86.3	180207	12	AC060802	AC060802 Homo sapi

238	16.4	86.3	180339	12	AC109839	AC109839 Rattus no	311	16.4	86.3	218992	12	AC103903	AC103903 Canis fam
c 239	16.4	86.3	180638	5	AC068763	AC068763 Homo sapi	312	16.4	86.3	222212	12	AC149709	AC149709 Bos tauru
240	16.4	86.3	180967	6	AC135289	AC135289 Mus muscu	313	16.4	86.3	222869	12	AC113777	AC113777 Rattus no
c 241	16.4	86.3	181895	12	AC117068	AC117068 Rattus no	314	16.4	86.3	223266	6	AC023174	AC023174 Mus muscu
242	16.4	86.3	182471	12	AC151133	AC151133 Bos tauru	c 315	16.4	86.3	223627	12	AC097891	AC097891 Rattus no
243	16.4	86.3	182533	12	AC176101	AC176101 Strongylo	c 316	16.4	86.3	223672	12	AC121720	AC121720 Rattus no
c 244	16.4	86.3	183248	12	AC161850	AC161850 Bos tauru	c 317	16.4	86.3	223936	12	AC103442	AC103442 Rattus no
c 245	16.4	86.3	183355	12	AC021801	AC021801 Homo sapi	c 318	16.4	86.3	223987	12	AC131556	AC131556 Rattus no
c 246	16.4	86.3	183596	5	AC147094	AC147094 Pan trogl	c 319	16.4	86.3	224198	12	AC099435	AC099435 Rattus no
c 247	16.4	86.3	184428	12	AC119645	AC119645 Rattus no	c 320	16.4	86.3	224691	12	AC129870	AC129870 Rattus no
c 248	16.4	86.3	184655	12	AC0979379	AC0979379 Homo sapi	c 321	16.4	86.3	224787	12	AC135144	AC135144 Rattus no
c 249	16.4	86.3	184666	12	AC013508	AC013508 Homo sapi	c 322	16.4	86.3	224809	12	AC163304	AC163304 Bos tauru
c 250	16.4	86.3	185174	5	AC073091	AC073091 Homo sapi	c 323	16.4	86.3	225083	6	AL596382	AL596382 Mouse DNA
251	16.4	86.3	186002	12	AC171251	AC171251 Bos tauru	324	16.4	86.3	225600	6	AC115766	AC115766 Mus muscu
c 252	16.4	86.3	186078	5	AC037482	AC037482 Homo sapi	325	16.4	86.3	225600	6	AC095532	AC095532 Rattus no
c 253	16.4	86.3	186302	12	AC160700	AC160700 Bos tauru	326	16.4	86.3	225756	6	AC166390	AC166390 Bos tauru
c 254	16.4	86.3	186737	12	AC093207	AC093207 Homo sapi	c 327	16.4	86.3	229427	12	AC136855	AC136855 Rattus no
c 255	16.4	86.3	187416	6	AC121959	AC121959 Mus muscu	c 328	16.4	86.3	229932	12	AC136855	AC136855 Rattus no
c 256	16.4	86.3	188965	12	AC173035	AC173035 Bos tauru	c 329	16.4	86.3	230352	12	AC016982	AC016982 Mus muscu
c 257	16.4	86.3	189130	6	AC132110	AC132110 Mus muscu	c 330	16.4	86.3	230696	5	AC007221	AC007221 Homo sapi
c 258	16.4	86.3	189245	12	AC068090	AC068090 Homo sapi	c 331	16.4	86.3	231600	12	AC130779	AC130779 Rattus no
c 259	16.4	86.3	189505	6	AC134540	AC134540 Mus muscu	c 332	16.4	86.3	233443	12	AC097704	AC097704 Rattus no
c 260	16.4	86.3	189536	5	AC117834	AC117834 Homo sapi	c 333	16.4	86.3	234506	12	AC129866	AC129866 Rattus no
c 261	16.4	86.3	190631	12	AC011842	AC011842 Homo sapi	c 334	16.4	86.3	237344	12	AC096256	AC096256 Rattus no
c 262	16.4	86.3	190088	6	AC132446	AC132446 Mus muscu	c 335	16.4	86.3	238744	12	AC170522	AC170522 Bos tauru
c 263	16.4	86.3	190453	6	AC157919	AC157919 Mus muscu	c 336	16.4	86.3	239022	12	AC105496	AC105496 Rattus no
c 264	16.4	86.3	190775	12	CT573218	CT573218 Macaca mu	c 337	16.4	86.3	239655	12	AC096969	AC096969 Rattus no
c 265	16.4	86.3	191014	6	AC118540	AC118540 Mus muscu	c 338	16.4	86.3	242353	12	AC110824	AC110824 Rattus no
c 266	16.4	86.3	191119	6	AC172106	AC172106 Mus muscu	c 339	16.4	86.3	243197	12	AC166581	AC166581 Bos tauru
c 267	16.4	86.3	191606	5	AL162376	AL162376 Human DNA	c 340	16.4	86.3	243409	12	AC098607	AC098607 Rattus no
c 268	16.4	86.3	192074	5	AC084758	AC084758 Homo sapi	341	16.4	86.3	245449	12	AC131606	AC131606 Rattus no
c 269	16.4	86.3	192888	12	AC172554	AC172554 Bos tauru	c 342	16.4	86.3	247108	12	AC123333	AC123333 Rattus no
c 270	16.4	86.3	192983	12	AC074122	AC074122 Homo sapi	c 343	16.4	86.3	249211	12	AC112539	AC112539 Rattus no
c 271	16.4	86.3	193211	6	AC130823	AC130823 Mus muscu	c 344	16.4	86.3	249291	12	AC109718	AC109718 Rattus no
c 272	16.4	86.3	193220	5	AC007218	AC007218 Homo sapi	c 345	16.4	86.3	249990	12	AC094813	AC094813 Rattus no
c 273	16.4	86.3	193519	5	AC060231	AC060231 Homo sapi	c 346	16.4	86.3	251452	12	AC097208	AC097208 Rattus no
c 274	16.4	86.3	193948	12	AC172834	AC172834 Bos tauru	c 347	16.4	86.3	251891	12	AC084742	AC084742 Rattus no
c 275	16.4	86.3	194221	6	AC113998	AC113998 Mus muscu	c 348	16.4	86.3	253880	12	AC113716	AC113716 Rattus no
c 276	16.4	86.3	194350	12	AC128084	AC128084 Rattus no	c 349	16.4	86.3	254264	12	AC128961	AC128961 Rattus no
c 277	16.4	86.3	195429	6	AC122241	AC122241 Mus muscu	c 350	16.4	86.3	254329	12	AC121061	AC121061 Rattus no
c 278	16.4	86.3	196236	6	AC108780	AC108780 Mus muscu	c 351	16.4	86.3	255398	12	AC112553	AC112553 Rattus no
c 279	16.4	86.3	197925	6	AL672241	AL672241 Mouse DNA	c 352	16.4	86.3	259147	12	AC102975	AC102975 Rattus no
c 280	16.4	86.3	199642	12	AC155185	AC155185 Bos tauru	c 353	16.4	86.3	259798	12	AC103131	AC103131 Rattus no
c 281	16.4	86.3	199656	6	AC130661	AC130661 Mus muscu	c 354	16.4	86.3	262352	12	AC152313	AC152313 Bos tauru
c 282	16.4	86.3	200505	6	AC164560	AC164560 Mus muscu	c 355	16.4	86.3	262352	12	AC095905	AC095905 Rattus no
c 283	16.4	86.3	200540	6	AC152820	AC152820 Mus muscu	c 356	16.4	86.3	263970	12	AC095905	AC095905 Rattus no
c 284	16.4	86.3	200912	6	AC155162	AC155162 Mus muscu	c 357	16.4	86.3	266925	12	AC126841	AC126841 Rattus no
c 285	16.4	86.3	201194	12	AC139119	AC139119 Didelphis	c 358	16.4	86.3	267055	12	AC145871	AC145871 Pan trogl
c 286	16.4	86.3	201303	5	AC078851	AC078851 Homo sapi	c 359	16.4	86.3	270086	12	AC128132	AC128132 Rattus no
c 287	16.4	86.3	201416	12	AC130437	AC130437 Homo sapi	c 360	16.4	86.3	271961	12	AC107572	AC107572 Rattus no
c 288	16.4	86.3	201641	12	AC148904	AC148904 Oryzomys	c 361	16.4	86.3	271961	12	AC107572	AC107572 Rattus no
c 289	16.4	86.3	205419	12	AC165537	AC165537 Bos tauru	c 362	16.4	86.3	272677	12	AC097289	AC097289 Rattus no
c 290	16.4	86.3	208131	12	AC161827	AC161827 Mus muscu	c 363	16.4	86.3	272714	12	AC129447	AC129447 Rattus no
c 291	16.4	86.3	208396	6	AC102672	AC102672 Mus muscu	c 364	16.4	86.3	274928	12	AC114717	AC114717 Rattus no
c 292	16.4	86.3	209175	6	AL450321	AL450321 Mouse DNA	c 365	16.4	86.3	276378	12	AC096069	AC096069 Rattus no
c 293	16.4	86.3	209384	12	AC106620	AC106620 Rattus no	c 366	16.4	86.3	277588	12	AC164843	AC164843 Bos tauru
c 294	16.4	86.3	209823	12	AC112340	AC112340 Rattus no	c 367	16.4	86.3	277752	12	AC123001	AC123001 Rattus no
c 295	16.4	86.3	210014	6	AC127283	AC127283 Mus muscu	c 368	16.4	86.3	280364	12	AC098343	AC098343 Rattus no
c 296	16.4	86.3	210179	13	AC150414	AC150414 Branchios	c 369	16.4	86.3	281915	12	AC111493	AC111493 Rattus no
c 297	16.4	86.3	210385	12	AC009627	AC009627 Homo sapi	c 370	16.4	86.3	287768	12	AC130177	AC130177 Rattus no
c 298	16.4	86.3	210512	12	AC166818	AC166818 Mus muscu	c 371	16.4	86.3	293381	12	AC091340	AC091340 Rattus no
c 299	16.4	86.3	210851	6	AC117686	AC117686 Mus muscu	c 372	16.4	86.3	301225	12	AC111619	AC111619 Rattus no
c 300	16.4	86.3	210954	11	AL627213	AL627213 Gallus ga	c 373	16.4	86.3	306003	12	AC110456	AC110456 Rattus no
c 301	16.4	86.3	211343	6	AC152953	AC152953 Mus muscu	c 374	16.4	86.3	306349	12	AC145085	AC145085 Pan trogl
c 302	16.4	86.3	212206	6	AC168881	AC168881 Mus muscu	c 375	16.4	86.3	313183	12	AC095152	AC095152 Rattus no
c 303	16.4	86.3	212563	6	AC125146	AC125146 Mus muscu	c 376	16.4	86.3	315463	12	AC122070	AC122070 Rattus no
c 304	16.4	86.3	213411	12	AC111830	AC111830 Rattus no	c 377	16.4	86.3	330743	10	PBU42580	PBU42580 Paramedum
c 305	16.4	86.3	213411	12	AC095719	AC095719 Mus muscu	c 378	16.4	86.3	334431	12	AC117917	AC117917 Rattus no
c 306	16.4	86.3	215916	6	AL662922	AL662922 Mouse DNA	c 379	16.4	86.3	345682	12	AC095142	AC095142 Rattus no
c 307	16.4	86.3	217836	5	AC149231	AC149231 Pan trogl	c 380	16.4	84.2	430	13	AF108128	AF108128 Laemigona
c 308	16.4	86.3	217898	12	AC095909	AC095909 Rattus no	c 381	16.4	84.2	910	2	BD021579	BD021579 Novel gen
c 309	16.4	86.3	217991	12	AC130504	AC130504 Rattus no	c 382	16.4	84.2	981	2	BD021578	BD021578 Novel gen
c 310	16.4	86.3	218859	12	AC022912	AC022912 Homo sapi	c 383	16.4	84.2	981	2	BD101516	BD101516 Novel gen

C 384	16	84.2	2370	13	THEHYPOC	L48612 Theileria p	C 457	16	84.2	246582	12	AC171149	AC171149 Microcebu
C 385	16	84.2	3024	2	CO721988	CO721988 Sequence	C 458	16	84.2	258038	12	AC16018	AC16018 Mus muscu
C 386	16	84.2	3393	5	BT009924	BT009924 Homo sapi	C 459	16	84.2	262347	12	CR538721	CR538721 Danio rer
C 387	16	84.2	3393	8	AY889161	AY889161 Synthetic	C 460	16	84.2	266344	5	AC005158	AC005158 Homo sapi
C 388	16	84.2	3393	8	AY891700	AY891700 Synthetic	C 461	16	84.2	267735	12	AC126617	AC126617 Rattus no
C 389	16	84.2	3393	8	AY891701	AY891701 Synthetic	C 462	16	84.2	267735	12	AC120719	AC120719 Rattus no
C 390	16	84.2	3393	8	BT007570	BT007570 Synthetic	C 463	16	84.2	275277	12	AC098018	AC098018 Rattus no
C 391	16	84.2	3447	5	HSMB01233	HSMB01233	C 464	16	84.2	280810	12	AC046188	AC046188 Homo sapi
C 392	16	84.2	3737	2	DD165362	DD165362 NUCLEIC A	C 465	16	84.2	283075	6	AC107801	AC107801 Mus muscu
C 393	16	84.2	3784	5	BC013045	BC013045 Homo sapi	C 466	16	84.2	287119	12	AC161144	AC161144 Bos tauru
C 394	16	84.2	3815	5	AB209293	AB209293 Homo sapi	C 467	16	84.2	287862	12	AC155041	AC155041 Bos tauru
C 395	16	84.2	4022	5	HSU66616	HSU66616 Human SWI/S	C 468	16	84.2	304570	12	AC167052	AC167052 Bos tauru
C 396	16	84.2	4032	5	BC026222	BC026222 Homo sapi	C 469	16	84.2	349688	15	CR378671	CR378671 Photobact
C 397	16	84.2	4271	5	AB209006	AB209006 Homo sapi	C 470	15.8	83.2	546	7	BV328779	BV328779 S241P671F
C 398	16	84.2	6079	15	AB219543	AB219543 Campyloba	C 471	15.8	83.2	582	7	BV326885	BV326885 S241P6223
C 399	16	84.2	6452	15	AB219544	AB219544 Campyloba	C 472	15.8	83.2	590	7	BV351834	BV351834 S230P6352
C 400	16	84.2	62896	12	AC120151	AC120151 Mus muscu	C 473	15.8	83.2	593	7	G81969	G81969 S208P6706FB
C 401	16	84.2	66806	12	AC090185	AC090185 Homo sapi	C 474	15.8	83.2	611	7	G89434	G89434 S209P6068RD
C 402	16	84.2	61667	12	AC091270	AC091270 Homo sapi	C 475	15.8	83.2	615	7	BV351968	BV351968 S230P6352
C 403	16	84.2	83714	5	AL136125	AL136125 Human DNA	C 476	15.8	83.2	640	7	BV291181	BV291181 S232P6206
C 404	16	84.2	88761	12	AC156059	AC156059 Bos tauru	C 477	15.8	83.2	644	7	BV038993	BV038993 S212P6048
C 405	16	84.2	96564	11	CR762431	CR762431 Zebrafish	C 478	15.8	83.2	670	7	BV483698	BV483698 S215P6835
C 406	16	84.2	97476	11	BX470121	BX470121 Zebrafish	C 479	15.8	83.2	701	7	BV474563	BV474563 G591P6270
C 407	16	84.2	101956	12	AP002494	AP002494 Homo sapi	C 480	15.8	83.2	740	7	BV532206	BV532206 G591P6232
C 408	16	84.2	102873	4	ATF26G5	ATF26G5	C 481	15.8	83.2	778	11	CT0255345	CT0255345 Xenopus t
C 409	16	84.2	103391	2	CS086349	CS086349 Sequence	C 482	15.8	83.2	806	7	BV019839	BV019839 S212P6027
C 410	16	84.2	106383	5	AC093753	AC093753 Homo sapi	C 483	15.8	83.2	844	7	BV511189	BV511189 qbk90A03.
C 411	16	84.2	110000	12	AC128438_2	Continuation (3 of	C 484	15.8	83.2	854	7	BV547208	BV547208 rtm52D06.
C 412	16	84.2	111079	5	HS419C19	HS419C19 Human DNA	C 485	15.8	83.2	965	7	BV646129	BV646129 S217P6092
C 413	16	84.2	110282	12	AC078973	AC078973 Homo sapi	C 486	15.8	83.2	1056	4	AF407335	AF407335 Lentinula
C 414	16	84.2	138978	11	CR388214	CR388214 Zebrafish	C 487	15.8	83.2	1131	2	CQ804518	CQ804518 Sequence
C 415	16	84.2	142525	5	AL105752	AL105752 Homo sapi	C 488	15.8	83.2	1366	2	AX700530	AX700530 Sequence
C 416	16	84.2	144844	6	AL845417	AL845417 Mouse DNA	C 489	15.8	83.2	1398	4	AY087822	AY087822 Arabidops
C 417	16	84.2	146596	5	AP000497	AP000497 Homo sapi	C 490	15.8	83.2	1515	11	CR386597	CR386597 Gallus ga
C 418	16	84.2	152224	5	AC144536	AC144536 Homo sapi	C 491	15.8	83.2	1599	2	DD196812	DD196812 HUMAN GEN
C 419	16	84.2	158431	12	CT030707	CT030707 Danio rer	C 492	15.8	83.2	1752	2	AR507118	AR507118 Sequence
C 420	16	84.2	159324	5	AC146165	AC146165 Pan trogl	C 493	15.8	83.2	1797	11	BC055179	BC055179 Danio rer
C 421	16	84.2	160390	12	AC068186	AC068186 Homo sapi	C 494	15.8	83.2	1868	2	AX700531	AX700531 Sequence
C 422	16	84.2	161949	6	CT485613	CT485613 Mouse DNA	C 495	15.8	83.2	1987	11	CR848091	CR848091 Xenopus t
C 423	16	84.2	163196	5	AC005220	AC005220 Homo sapi	C 496	15.8	83.2	2063	2	BD158476	BD158476 Primer fo
C 424	16	84.2	167380	5	AP006242	AP006242 Homo sapi	C 497	15.8	83.2	2063	2	AX880602	AX880602 Sequence
C 425	16	84.2	189031	6	AC124476	AC124476 Mus muscu	C 498	15.8	83.2	2063	5	AX023467	AX023467 Homo sapi
C 426	16	84.2	170393	12	AC024304	AC024304 Homo sapi	C 499	15.8	83.2	2094	2	CQ600300	CQ600300 Sequence
C 427	16	84.2	172866	12	AC164240	AC164240 Bos tauru	C 500	15.8	83.2	2141	5	AB179345	AB179345 Macaca fa
C 428	16	84.2	173744	6	AC132279	AC132279 Mus muscu	C 501	15.8	83.2	2165	2	CQ577719	CQ577719 Sequence
C 429	16	84.2	174474	5	AL591408	AL591408 Human DNA	C 502	15.8	83.2	2175	2	AX430350	AX430350 Sequence
C 430	16	84.2	174487	5	AC073896	AC073896 Homo sapi	C 503	15.8	83.2	2208	2	CS073462	CS073462 Sequence
C 431	16	84.2	179186	12	AL357113	AL357113 Homo sapi	C 504	15.8	83.2	2208	2	CS073463	CS073463 Sequence
C 432	16	84.2	179282	12	AC173683	AC173683 Bos tauru	C 505	15.8	83.2	2208	2	CS073464	CS073464 Sequence
C 433	16	84.2	179659	12	AC122121	AC122121 Mus muscu	C 506	15.8	83.2	2208	2	CS073465	CS073465 Sequence
C 434	16	84.2	181313	12	AC021963	AC021963 Homo sapi	C 507	15.8	83.2	2208	2	CS073466	CS073466 Sequence
C 435	16	84.2	184762	12	AC023500	AC023500 Homo sapi	C 508	15.8	83.2	2208	2	CS073467	CS073467 Sequence
C 436	16	84.2	187568	5	AP002967	AP002967 Homo sapi	C 509	15.8	83.2	2208	2	CS073468	CS073468 Sequence
C 437	16	84.2	197018	6	AL59188	AL59188 Mus muscu	C 510	15.8	83.2	2208	2	CS073469	CS073469 Sequence
C 438	16	84.2	198460	6	AL807808	AL807808 Mouse DNA	C 511	15.8	83.2	2208	2	CS073470	CS073470 Sequence
C 439	16	84.2	200997	6	AC155244	AC155244 Mus muscu	C 512	15.8	83.2	2208	2	CS073471	CS073471 Sequence
C 440	16	84.2	202584	6	AC159321	AC159321 Mus muscu	C 513	15.8	83.2	2208	2	CS073480	CS073480 Sequence
C 441	16	84.2	204042	6	AC132456	AC132456 Mus muscu	C 514	15.8	83.2	2208	2	CS073481	CS073481 Sequence
C 442	16	84.2	205379	6	AL591174	AL591174 Mouse DNA	C 515	15.8	83.2	2208	2	CS073482	CS073482 Sequence
C 443	16	84.2	208809	11	BX927218	BX927218 Zebrafish	C 516	15.8	83.2	2208	10	AY530576	AY530576 Adeno-ass
C 444	16	84.2	208817	6	AC103934	AC103934 Mus muscu	C 517	15.8	83.2	2208	10	AY530577	AY530577 Adeno-ass
C 445	16	84.2	209081	12	AC150400	AC150400 Branchios	C 518	15.8	83.2	2208	10	AY530580	AY530580 Adeno-ass
C 446	16	84.2	209319	12	AC107799	AC107799 Mus muscu	C 519	15.8	83.2	2208	10	AY530581	AY530581 Adeno-ass
C 447	16	84.2	213854	12	AC117001	AC117001 Rattus no	C 520	15.8	83.2	2208	10	AY530583	AY530583 Adeno-ass
C 448	16	84.2	216945	12	AC1156619	AC1156619 Mus muscu	C 521	15.8	83.2	2208	10	AY530593	AY530593 Adeno-ass
C 449	16	84.2	218283	12	AC079476	AC079476 Mus muscu	C 522	15.8	83.2	2208	10	AY530594	AY530594 Adeno-ass
C 450	16	84.2	225709	12	AC096026	AC096026 Rattus no	C 523	15.8	83.2	2208	10	AY530628	AY530628 Adeno-ass
C 451	16	84.2	228010	12	CR533429	CR533429 Danio rer	C 524	15.8	83.2	2246	11	CR848328	CR848328 Xenopus t
C 452	16	84.2	232182	12	AC112049	AC112049 Rattus no	C 525	15.8	83.2	2260	5	AF295096	AF295096 Homo sapi
C 453	16	84.2	239246	12	AC102997	AC102997 Rattus no	C 526	15.8	83.2	2295	5	BC009921	BC009921 Homo sapi
C 454	16	84.2	241519	12	AC112461	AC112461 Rattus no	C 527	15.8	83.2	2295	5	BC009921	BC009921 Homo sapi
C 455	16	84.2	241953	12	AC162843	AC162843 Bos tauru	C 528	15.8	83.2	2305	11	CR382437	CR382437 Gallus ga
C 456	16	84.2	241953	12	AC162843	AC162843 Bos tauru	C 529	15.8	83.2	2328	11	BX950379	BX950379 Gallus ga

C 530	15.8	83.2	2329	5	AF309561	AF309561 Homo sapi	603	15.8	83.2	98512	11	BX005060	BX005060 Zebrafish
C 531	15.8	83.2	2376	2	AX700518	AX700518 Sequence	604	15.8	83.2	100000	5	AB020875	AB020875 Homo sapi
C 532	15.8	83.2	2412	2	AR339532	AR339532 Sequence	C 605	15.8	83.2	100547	5	AC008637	AC008637 Homo sapi
C 533	15.8	83.2	2438	2	CQ492700	CQ492700 Sequence	C 606	15.8	83.2	101324	5	AC010386	AC010386 Homo sapi
C 534	15.8	83.2	2438	2	CQ497329	CQ497329 Sequence	C 607	15.8	83.2	102230	5	HS23445	HS23445 Homo sapi
C 535	15.8	83.2	2441	2	CQ596057	CQ596057 Sequence	608	15.8	83.2	102928	12	AC178880	AC178880 Strongylo
C 536	15.8	83.2	2459	2	AX427519	AX427519 Sequence	609	15.8	83.2	103244	14	CR956378	CR956378 Pig DNA s
C 537	15.8	83.2	2833	2	AR510319	AR510319 Sequence	610	15.8	83.2	103488	5	AL805915	AL805915 Human DNA
C 538	15.8	83.2	3111	13	BT006012	BT006012 Drosophil	611	15.8	83.2	104079	12	AY382680	AY382680 Braasica
C 539	15.8	83.2	4329	6	BC057374	BC057374 Mus muscu	C 612	15.8	83.2	104542	5	AF000566	AF000566 Homo sapi
C 540	15.8	83.2	4331	6	AK122560	AK122560 Mus muscu	613	15.8	83.2	105546	12	AC175954	AC175954 Strongylo
C 541	15.8	83.2	4374	6	BC055330	BC055330 Mus muscu	C 614	15.8	83.2	107330	5	AC019178	AC019178 Homo sapi
C 542	15.8	83.2	4385	6	BC043127	BC043127 Mus muscu	615	15.8	83.2	108556	12	AC170551	AC170551 Bos tauru
C 543	15.8	83.2	4590	2	CQ842369	CQ842369 Sequence	C 616	15.8	83.2	108752	6	AF084363	AF084363 Mus muscu
C 544	15.8	83.2	4590	5	AK125326	AK125326 Homo sapi	617	15.8	83.2	108709	12	AC178156	AC178156 Strongylo
C 545	15.8	83.2	5935	11	BC073499	BC073499 Xenopus l	C 618	15.8	83.2	110000	2	BD430793	BD430793
C 546	15.8	83.2	8058	4	AB111462	AB111462 Coprinops	619	15.8	83.2	110000	4	AF008212	AF008212
C 547	15.8	83.2	9640	5	AC093375	AC093375 Homo sapi	620	15.8	83.2	110000	4	AF008213	AF008213
C 548	15.8	83.2	10807	15	AE013148	AE013148 Thermoana	621	15.8	83.2	110000	4	AF008214	AF008214
C 549	15.8	83.2	11682	2	CQ731076	CQ731076 Sequence	C 622	15.8	83.2	110000	4	AF008215	AF008215
C 550	15.8	83.2	12462	2	CQ577718	CQ577718 Sequence	C 623	15.8	83.2	110000	4	AF008215	AF008215
C 551	15.8	83.2	12481	2	CQ600299	CQ600299 Sequence	C 624	15.8	83.2	110000	4	AF008218	AF008218
C 552	15.8	83.2	19664	15	AF343837S1	AF343837 Enterococ	626	15.8	83.2	110000	4	AF008218	AF008218
C 553	15.8	83.2	26205	11	AL713959	AL713959 Zebrafish	C 627	15.8	83.2	110000	4	AF008218	AF008218
C 554	15.8	83.2	36116	5	AC078982	AC078982 Homo sapi	628	15.8	83.2	110000	4	CR382130	CR382130
C 555	15.8	83.2	37897	12	AC162773	AC162773 Macropus	C 629	15.8	83.2	110000	4	AE017352	AE017352
C 556	15.8	83.2	37971	6	AC170995	AC170995 Mus muscu	C 630	15.8	83.2	110000	4	AF008207	AF008207
C 557	15.8	83.2	40446	5	AC005925	AC005925 Homo sapi	631	15.8	83.2	110000	4	AF008207	AF008207
C 558	15.8	83.2	45300	12	AC136350	AC136350 Homo sapi	C 632	15.8	83.2	110000	4	AF008211	AF008211
C 559	15.8	83.2	47530	12	AC101037	AC101037 Mus muscu	C 633	15.8	83.2	110000	4	AF008211	AF008211
C 560	15.8	83.2	48012	5	AY6112854	AY6112854 Homo sapi	C 634	15.8	83.2	110000	5	DQ354388	DQ354388
C 561	15.8	83.2	49218	5	AL391474	AL391474 Human DNA	C 635	15.8	83.2	110000	5	DQ354389	DQ354389
C 562	15.8	83.2	50978	13	AC002445	AC002445 Drosophil	C 636	15.8	83.2	110000	5	DQ354390	DQ354390
C 563	15.8	83.2	52622	12	AC100195	AC100195 Mus muscu	C 637	15.8	83.2	110000	5	DQ354391	DQ354391
C 564	15.8	83.2	56230	12	AC023664	AC023664 Homo sapi	C 638	15.8	83.2	110000	5	DQ356257	DQ356257
C 565	15.8	83.2	57981	11	AL731850	AL731850 Zebrafish	C 639	15.8	83.2	110000	5	DQ356258	DQ356258
C 566	15.8	83.2	58815	5	AC079405	AC079405 Homo sapi	C 640	15.8	83.2	110000	5	DQ356259	DQ356259
C 567	15.8	83.2	59040	12	AC022009	AC022009 Homo sapi	C 641	15.8	83.2	110000	5	DQ356260	DQ356260
C 568	15.8	83.2	60232	5	AC073902	AC073902 Homo sapi	C 642	15.8	83.2	110000	5	DQ356261	DQ356261
C 569	15.8	83.2	60375	12	AC174818	AC174818 Bos tauru	C 643	15.8	83.2	110000	5	DQ356262	DQ356262
C 570	15.8	83.2	66128	12	BX679661	Continuation (5 of	C 644	15.8	83.2	110000	5	DQ356263	DQ356263
C 571	15.8	83.2	66367	12	AC136298	AC136298 Homo sapi	C 645	15.8	83.2	110000	5	DQ356264	DQ356264
C 572	15.8	83.2	66657	12	AC118049	AC118049 Mus muscu	C 646	15.8	83.2	110000	12	AC091239	AC091239
C 573	15.8	83.2	68229	5	AB119283	AB119283 Homo sapi	C 647	15.8	83.2	110000	12	AC091242	AC091242
C 574	15.8	83.2	69074	12	AC020360	AC020360 Drosophil	C 648	15.8	83.2	110000	12	AC091347	AC091347
C 575	15.8	83.2	70607	6	AT038861S16	AT038867 Mus muscu	C 649	15.8	83.2	110000	12	AC111637	AC111637
C 576	15.8	83.2	73481	5	AC011891	AC011891 Homo sapi	C 650	15.8	83.2	110000	12	AC111637	AC111637
C 577	15.8	83.2	75247	12	AC165562	AC165562 Bos tauru	C 651	15.8	83.2	110000	12	AC114343	AC114343
C 578	15.8	83.2	77463	11	BX950206	BX950206 Zebrafish	C 652	15.8	83.2	110000	12	AC123221	AC123221
C 579	15.8	83.2	78318	5	AP004146	AP004146 Homo sapi	C 653	15.8	83.2	110000	12	AL845451	AL845451
C 580	15.8	83.2	78474	12	AC176200	AC176200 Strongylo	C 654	15.8	83.2	110000	12	CT572984	CT572984
C 581	15.8	83.2	79394	12	AC016339	AC016339 Homo sapi	C 655	15.8	83.2	110000	15	CP000099	CP000099
C 582	15.8	83.2	81731	12	AP003529	AP003529 Oryza sat	C 656	15.8	83.2	110000	15	CP000099	CP000099
C 583	15.8	83.2	8336	4	AB019235	AB019235 Arabidops	C 657	15.8	83.2	110000	15	CP000119	CP000119
C 584	15.8	83.2	83268	12	AC167093	AC167093 Sorex ara	C 658	15.8	83.2	110000	15	AE008384	AE008384
C 585	15.8	83.2	84514	6	AL627264	AL627264 Mouse DNA	C 659	15.8	83.2	110000	15	AE015928	AE015928
C 586	15.8	83.2	85144	5	AC125621	AC125621 Homo sapi	C 660	15.8	83.2	110000	15	AE016853	AE016853
C 587	15.8	83.2	85946	5	AC004240	AC004240 Homo sapi	C 661	15.8	83.2	110000	15	AE017199	AE017199
C 588	15.8	83.2	87394	4	AC005882	AC005882 Arabidops	C 662	15.8	83.2	110000	15	BA000011	BA000011
C 589	15.8	83.2	88502	12	AC019167	AC019167 Homo sapi	C 663	15.8	83.2	110250	12	AC134022	AC134022
C 590	15.8	83.2	89297	5	AC109996	AC109996 Homo sapi	C 664	15.8	83.2	110310	6	AF131866	AF131866
C 591	15.8	83.2	89521	5	HS7511L1	AL031675 Human DNA	C 665	15.8	83.2	114356	5	AC114934	AC114934
C 592	15.8	83.2	89885	13	AC004368	AC004368 Drosophil	C 666	15.8	83.2	115027	6	CT573035	CT573035
C 593	15.8	83.2	90408	12	AP008135	AP008135 Lotus cor	C 667	15.8	83.2	116171	12	AC181131	AC181131 Strongylo
C 594	15.8	83.2	90435	12	AC138794	AC138794 Homo sapi	C 668	15.8	83.2	119691	4	AP003743	AP003743 Oryza sat
C 595	15.8	83.2	90946	6	AC147595	AC147595 Mus muscu	C 669	15.8	83.2	119916	5	AF241734	AF241734 Homo sapi
C 596	15.8	83.2	91017	12	AC179226	AC179226 Strongylo	C 670	15.8	83.2	120978	5	AL732367	AL732367 Human DNA
C 597	15.8	83.2	92056	12	AC176182	AC176182 Strongylo	C 671	15.8	83.2	121538	6	AC129295	AC129295 Mus muscu
C 598	15.8	83.2	92998	5	AL731546	AL731546 Human DNA	C 672	15.8	83.2	122492	4	AC149131	AC149131 Medicago
C 599	15.8	83.2	95449	5	AC148208	AC148208 Callicebu	C 673	15.8	83.2	123423	6	AC006949	AC006949 Mus muscu
C 600	15.8	83.2	96334	12	AC021343	AC021343 Homo sapi	C 674	15.8	83.2	123521	12	AC180267	AC180267 Strongylo
C 601	15.8	83.2	96628	11	AC156247	AC156247 Xenopus t	C 675	15.8	83.2	123547	12	AC015857	AC015857 Homo sapi
C 602	15.8	83.2	97681	11	AC156343	AC156343 Xenopus t							

676	15.8	83.2	124250	5	AC093387	AC093387 Homo sapi	749	15.8	83.2	154539	12	AC079187	AC079187 Homo sapi
677	15.8	83.2	124959	4	CNS071PQ	AL713947 Oryza sat	750	15.8	83.2	154824	5	AC138917	AC138917 Homo sapi
c 678	15.8	83.2	125482	5	AC108924	AC108924 Homo sapi	c 751	15.8	83.2	154867	5	AC017035	AC017035 Homo sapi
679	15.8	83.2	126123	11	AL929116	AL929116 Zebrafish	c 752	15.8	83.2	155000	5	AP006293	AP006293 Homo sapi
680	15.8	83.2	127678	2	CQ870038	CQ870038 Sequence	c 753	15.8	83.2	155122	12	AC141847	AC141847 Bos taurus
c 681	15.8	83.2	129081	12	AC163247	AC163247 Loxodonta	c 754	15.8	83.2	155397	12	AC148430	AC148430 Bos taurus
c 682	15.8	83.2	129218	12	AC138851	AC138851 Homo sapi	755	15.8	83.2	156102	12	AC145182	AC145182 Macropus
c 683	15.8	83.2	129949	12	AC138851	AC138851 Homo sapi	c 756	15.8	83.2	156503	11	BX664752	BX664752 Zebrafish
c 684	15.8	83.2	130949	5	AC022940	AC022940 Homo sapi	c 757	15.8	83.2	156543	14	AC098687	AC098687 Bos taurus
c 685	15.8	83.2	131011	12	AL731777	AL731777 Human DNA	c 758	15.8	83.2	156575	5	AC144587	AC144587 Homo sapi
c 686	15.8	83.2	131094	4	AC138087	AC138087 Medicago	759	15.8	83.2	156709	5	AB020863	AB020863 Homo sapi
c 687	15.8	83.2	131826	5	AC011460	AC011460 Homo sapi	c 760	15.8	83.2	156909	5	AC018639	AC018639 Human Chr
688	15.8	83.2	132028	11	EX890558	EX890558 Zebrafish	c 761	15.8	83.2	157579	5	AC018639	AC018639 Human Chr
689	15.8	83.2	132099	4	AC166741	AC166741 Brassica	c 762	15.8	83.2	158293	12	AL019706	AL019706 Drosophil
690	15.8	83.2	133301	4	AP003198	AP003198 Oryza sat	c 763	15.8	83.2	158310	5	AL162571	AL162571 Human DNA
691	15.8	83.2	136156	12	AC144358	AC144358 Pan trogl	c 764	15.8	83.2	158350	11	AL935163	AL935163 Zebrafish
c 692	15.8	83.2	136333	4	AC131248	AC131248 Medicago	765	15.8	83.2	159082	5	AP000940	AP000940 Homo sapi
c 693	15.8	83.2	136858	4	AC134345	AC134345 Oryza sat	766	15.8	83.2	159231	5	AL161654	AL161654 Human DNA
c 694	15.8	83.2	136880	12	AC078787	AC078787 Homo sapi	767	15.8	83.2	159840	5	AC151966	AC151966 Mus muscu
c 695	15.8	83.2	137077	5	AC007562	AC007562 Homo sapi	c 768	15.8	83.2	160392	5	AC104236	AC104236 Homo sapi
c 696	15.8	83.2	137191	12	AC180801	AC180801 Strongylo	769	15.8	83.2	160977	12	AC140906	AC140906 Homo sapi
c 697	15.8	83.2	137682	5	AC023274	AC023274 Homo sapi	c 770	15.8	83.2	161224	12	AC175724	AC175724 Bos tauru
c 698	15.8	83.2	138251	5	AC004908	AC004908 Homo sapi	c 771	15.8	83.2	161695	12	AC103901	AC103901 Canis fam
c 699	15.8	83.2	138363	12	AC168512	AC168512 Strongylo	c 772	15.8	83.2	161725	12	AC178307	AC178307 Strongylo
700	15.8	83.2	138369	12	AC079546	AC079546 Mus muscu	c 773	15.8	83.2	161893	5	AC137781	AC137781 Homo sapi
701	15.8	83.2	138859	5	AL359076	AL359076 Human DNA	c 774	15.8	83.2	162112	12	AC015265	AC015265 Mus muscu
702	15.8	83.2	139589	5	AC010470	AC010470 Homo sapi	c 775	15.8	83.2	162151	6	AC102841	AC102841 Mus muscu
c 703	15.8	83.2	140221	12	AC132495	AC132495 Oryza sat	c 776	15.8	83.2	162162	12	AC013294	AC013294 Homo sapi
c 704	15.8	83.2	140269	12	AC176271	AC176271 Strongylo	c 777	15.8	83.2	162454	6	BX470216	BX470216 Mouse Chr
c 705	15.8	83.2	140544	12	AC011798	AC011798 Homo sapi	c 778	15.8	83.2	162691	6	AC003061	AC003061 Mouse Chr
c 706	15.8	83.2	140544	12	AC011798	AC011798 Homo sapi	c 779	15.8	83.2	163542	11	BX571694	BX571694 Zebrafish
c 707	15.8	83.2	140936	12	AC155366	AC155366 zea mays	c 780	15.8	83.2	163954	12	AC017155	AC017155 Drosophil
c 708	15.8	83.2	141007	12	AC025764	AC025764 Homo sapi	c 781	15.8	83.2	164689	12	AL953873	AL953873 Homo sapi
c 709	15.8	83.2	141854	12	AC021595	AC021595 Homo sapi	c 782	15.8	83.2	164893	6	AC084382	AC084382 Mus muscu
c 710	15.8	83.2	142235	12	AP005973	AP005973 Oryza sat	c 783	15.8	83.2	164936	6	AC156393	AC156393 Mus muscu
711	15.8	83.2	142670	4	CNS08C9U	AL772416 Oryza sat	c 784	15.8	83.2	165443	12	AC079939	AC079939 Mus muscu
c 712	15.8	83.2	142776	12	AP005109	AP005109 Oryza sat	c 785	15.8	83.2	165502	5	BS000051	BS000051 Pan trogl
c 713	15.8	83.2	143416	12	AC015470	AC015470 Homo sapi	c 786	15.8	83.2	165831	12	AC168567	AC168567 Strongylo
c 714	15.8	83.2	143517	5	AL356752	AL356752 Human DNA	c 787	15.8	83.2	165930	12	AC073509	AC073509 Homo sapi
c 715	15.8	83.2	143860	5	AC009155	AC009155 Homo sapi	c 788	15.8	83.2	166062	12	AC118449	AC118449 Rattus no
c 716	15.8	83.2	144694	4	AP005765	AP005765 Oryza sat	c 789	15.8	83.2	166980	12	AC152428	AC152428 Bos tauru
c 717	15.8	83.2	144709	12	AC120084	AC120084 Rattus no	c 790	15.8	83.2	167269	6	AC140291	AC140291 Mus muscu
c 718	15.8	83.2	144894	12	AC182362	AC182362 Spermothi	c 791	15.8	83.2	167269	6	AC121982	AC121982 Mus muscu
719	15.8	83.2	145156	5	AC010414	AC010414 Homo sapi	c 792	15.8	83.2	167409	5	AC121982	AC121982 Mus muscu
720	15.8	83.2	145274	12	AC019901	AC019901 Drosophil	c 793	15.8	83.2	167523	5	AC145983	AC145983 Pan trogl
721	15.8	83.2	145363	12	AC079213	AC079213 Homo sapi	c 794	15.8	83.2	167587	4	AP002914	AP002914 Oryza sat
722	15.8	83.2	145381	12	AC141341	AC141341 Rattus no	c 795	15.8	83.2	167589	12	AC176810	AC176810 Strongylo
c 723	15.8	83.2	145493	12	AC138839	AC138839 Homo sapi	c 796	15.8	83.2	167735	6	AC122818	AC122818 Mus muscu
724	15.8	83.2	146050	12	AC040943	AC040943 Homo sapi	c 797	15.8	83.2	167777	5	AC019330	AC019330 Homo sapi
725	15.8	83.2	146106	12	CR974486	CR974486 Mus muscu	c 798	15.8	83.2	168392	5	AC026461	AC026461 Homo sapi
726	15.8	83.2	146444	4	CNS08CC5	AL928775 Oryza sat	c 799	15.8	83.2	168460	12	AC011585	AC011585 Homo sapi
c 727	15.8	83.2	146961	12	AP007847	AP007847 Lotus cor	c 800	15.8	83.2	168515	6	AC161530	AC161530 Mus muscu
728	15.8	83.2	147079	5	AL772262	AL772262 Human DNA	c 801	15.8	83.2	168515	6	AC130215	AC130215 Mus muscu
729	15.8	83.2	147121	12	AC169910	AC169910 Macaca mu	c 802	15.8	83.2	168608	12	AL591477	AL591477 Homo sapi
730	15.8	83.2	147290	6	AC099608	AC099608 Mus muscu	c 803	15.8	83.2	168608	12	CR853296	CR853296 Dario rer
c 731	15.8	83.2	147811	11	CR847546	CR847546 Zebrafish	c 804	15.8	83.2	168830	12	AL591477	AL591477 Homo sapi
732	15.8	83.2	148349	12	AC182349	AC182349 Ateles ge	c 805	15.8	83.2	168912	6	AC024418	AC024418 Homo sapi
c 733	15.8	83.2	148488	12	AC167307	AC167307 Sorex ara	c 806	15.8	83.2	168981	12	AC126347	AC126347 Homo sapi
c 734	15.8	83.2	149435	4	AC136521	AC136521 Oryza sat	c 807	15.8	83.2	169199	5	AC026463	AC026463 Homo sapi
c 735	15.8	83.2	149505	5	AC093767	AC093767 Homo sapi	c 808	15.8	83.2	169209	6	AC102380	AC102380 Mus muscu
736	15.8	83.2	149585	12	AC146994	AC146994 Lytechinu	c 809	15.8	83.2	169319	12	AC182094	AC182094 Bos tauru
c 737	15.8	83.2	149675	12	AC140988	AC140988 Rattus no	c 810	15.8	83.2	169339	5	AC005082	AC005082 Homo sapi
c 738	15.8	83.2	150099	4	AC135916	AC135916 Genomic s	c 811	15.8	83.2	169856	5	AL591944	AL591944 Human DNA
c 739	15.8	83.2	152449	12	AC080175	AC080175 Homo sapi	c 812	15.8	83.2	169879	12	AL591944	AL591944 Homo sapi
740	15.8	83.2	153190	5	AC068597	AC068597 Homo sapi	c 813	15.8	83.2	170233	12	AC068844	AC068844 Homo sapi
741	15.8	83.2	153200	12	AC175827	AC175827 Sorex ara	c 814	15.8	83.2	170299	5	AC021523	AC021523 Homo sapi
742	15.8	83.2	153389	12	AC011695	AC011695 Homo sapi	c 815	15.8	83.2	170328	12	AC024169	AC024169 Homo sapi
743	15.8	83.2	153604	13	AC104509	AC104509 Drosophil	c 816	15.8	83.2	170350	5	AC068312	AC068312 Homo sapi
c 744	15.8	83.2	154065	12	AC084079	AC084079 Homo sapi	c 817	15.8	83.2	170755	5	AF176315	AF176315 Homo sapi
745	15.8	83.2	154160	12	AC046154	AC046154 Homo sapi	c 818	15.8	83.2	171075	12	AC026842	AC026842 Homo sapi
c 746	15.8	83.2	154348	5	AL691514	AL691514 Human DNA	c 819	15.8	83.2	171187	5	BS000062	BS000062 Pan trogl
747	15.8	83.2	154433	6	AL732447	AL732447 Mouse DNA	c 820	15.8	83.2	171206	5	AC090946	AC090946 Homo sapi
748	15.8	83.2	154474	4	AP006048	AP006048 Oryza sat	c 821	15.8	83.2	171360	12	AC068543	AC068543 Homo sapi

822	15.8	83.2	171376	12	AC171476	AC171476 Bos tauru	895	15.8	83.2	185297	6	AC102250	Mus muscu
823	15.8	83.2	171648	12	AC171403	AC171403 Mus muscu	896	15.8	83.2	185494	12	AC169055	Bos tauru
824	15.8	83.2	171716	5	AC034187	AC034187 Homo sapi	897	15.8	83.2	185617	6	AC140235	Mus muscu
825	15.8	83.2	171834	12	AC174213	AC174213 Bos tauru	898	15.8	83.2	186218	5	AP006248	Homo sapi
826	15.8	83.2	171937	5	AC008442	AC008442 Homo sapi	899	15.8	83.2	186311	6	AC166161	Mus muscu
827	15.8	83.2	172030	5	AC010254	AC010254 Homo sapi	900	15.8	83.2	186432	13	AC099034	Drosophil
828	15.8	83.2	172493	12	AC146888	AC146888 Bos tauru	901	15.8	83.2	186607	13	AC007772	Drosophil
829	15.8	83.2	172543	5	AP004287	AP004287 Homo sapi	902	15.8	83.2	186719	12	AC147574	Homo sapi
830	15.8	83.2	172579	5	AC008064	AC008064 Homo sapi	903	15.8	83.2	186747	12	AC018605	Homo sapi
831	15.8	83.2	172942	5	AC123537	AC123537 Macaca mu	904	15.8	83.2	187487	12	CT573220	Danio rer
832	15.8	83.2	173312	12	AC162865	AC162865 Mus muscu	905	15.8	83.2	187677	12	CT573220	Danio rer
833	15.8	83.2	173351	6	AC127562	AC127562 Mus muscu	906	15.8	83.2	187885	12	AC162985	Bos tauru
834	15.8	83.2	173534	5	AC007920	AC007920 Homo sapi	907	15.8	83.2	188483	12	AC176780	Strongylo
835	15.8	83.2	173819	12	AC161014	AC161014 Pan trogl	908	15.8	83.2	188665	12	CT025849	Danio rer
836	15.8	83.2	173821	5	AC020599	AC020599 Homo sapi	909	15.8	83.2	188864	12	AC135951	Macaca mu
837	15.8	83.2	174032	12	AC128150	AC128150 Rattus no	910	15.8	83.2	189033	6	AC117690	Mus muscu
838	15.8	83.2	174037	5	AC009995	AC009995 Homo sapi	911	15.8	83.2	189139	12	AC122947	Rattus no
839	15.8	83.2	174267	6	AC132327	AC132327 Mus muscu	912	15.8	83.2	189500	6	AC157656	Mus muscu
840	15.8	83.2	174333	5	AL356454	AL356454 Human DNA	913	15.8	83.2	189893	4	AP005298	Oryza sat
841	15.8	83.2	174407	5	AC138725	AC138725 Cercopith	914	15.8	83.2	189941	12	AC128854	Rattus no
842	15.8	83.2	174500	12	AC113089	AC113089 Mus muscu	915	15.8	83.2	190673	12	AC166022	Oryctolag
843	15.8	83.2	174507	6	AC132955	AC132955 Mus muscu	916	15.8	83.2	190679	12	AC147747	Homo sapi
844	15.8	83.2	174640	12	AC117840	AC117840 Rattus no	917	15.8	83.2	190680	5	AC147575	Homo sapi
845	15.8	83.2	174962	5	AC145145	AC145145 Homo sapi	918	15.8	83.2	190702	11	BX119962	Zebrafish
846	15.8	83.2	175291	12	AC177990	AC177990 Strongylo	919	15.8	83.2	190712	6	AC100382	Mus muscu
847	15.8	83.2	175706	12	AC013798	AC013798 Homo sapi	920	15.8	83.2	191462	12	AP005781	Oryza sat
848	15.8	83.2	175838	6	AC122495	AC122495 Mus muscu	921	15.8	83.2	191544	12	AC138855	Homo sapi
849	15.8	83.2	175992	5	AC016943	AC016943 Homo sapi	922	15.8	83.2	191662	6	AL772296	Mouse DNA
850	15.8	83.2	176018	5	AC093878	AC093878 Homo sapi	923	15.8	83.2	192026	6	AL576660	Mus muscu
851	15.8	83.2	176134	12	AC181356	AC181356 Strongylo	924	15.8	83.2	192044	5	AL590439	Human DNA
852	15.8	83.2	176222	11	AL929306	AL929306 Zebrafish	925	15.8	83.2	192356	5	AL621222	Colobus g
853	15.8	83.2	177191	5	AC145145	AC145145 Homo sapi	926	15.8	83.2	192499	12	AC170431	Bos tauru
854	15.8	83.2	177350	6	AC123743	AC123743 Mus muscu	927	15.8	83.2	192699	6	AC133597	Mus muscu
855	15.8	83.2	177535	12	AC109914	AC109914 Bos tauru	928	15.8	83.2	192944	12	AC146902	Callicebu
856	15.8	83.2	177917	6	AC115045	AC115045 Mus muscu	929	15.8	83.2	192961	5	AP002770	Homo sapi
857	15.8	83.2	178262	12	AC167516	AC167516 Sorex ara	930	15.8	83.2	193047	5	AC147069	Pan trogl
858	15.8	83.2	178273	6	AL954341	AL954341 Mouse DNA	931	15.8	83.2	193170	6	AC115709	Mus muscu
859	15.8	83.2	178605	12	AC160869	AC160869 Orolemur	932	15.8	83.2	193191	6	AL49276	Mus muscu
860	15.8	83.2	178781	6	AC154524	AC154524 Mus muscu	933	15.8	83.2	193219	6	AL807816	Mouse DNA
861	15.8	83.2	178965	6	AL844144	AL844144 Mouse DNA	934	15.8	83.2	193277	6	AC104324	Mus muscu
862	15.8	83.2	179125	13	AC009392	AC009392 Drosophil	935	15.8	83.2	193643	5	AC018644	Homo sapi
863	15.8	83.2	179476	11	BX294126	BX294126 Zebrafish	936	15.8	83.2	194020	12	AC101988	Mus muscu
864	15.8	83.2	179532	12	AC148894	AC148894 Orolemur	937	15.8	83.2	194095	5	AL929401	Human DNA
865	15.8	83.2	179723	5	AC087858	AC087858 Homo sapi	938	15.8	83.2	194439	12	AC109022	Rattus no
866	15.8	83.2	179777	12	AC067879	AC067879 Homo sapi	939	15.8	83.2	194476	12	AC147745	Homo sapi
867	15.8	83.2	180125	6	AC159186	AC159186 Mus muscu	940	15.8	83.2	194718	5	AP006254	Homo sapi
868	15.8	83.2	180152	5	BS000052	BS000052 Pan trogl	941	15.8	83.2	194748	12	BX890638	Mus muscu
869	15.8	83.2	180201	6	AC131736	AC131736 Mus muscu	942	15.8	83.2	195039	6	BX649621	Mouse DNA
870	15.8	83.2	180510	12	AC181368	AC181368 Strongylo	943	15.8	83.2	195169	5	AC017109	Homo sapi
871	15.8	83.2	180651	4	AP005745	AP005745 Oryza sat	944	15.8	83.2	195646	5	AC165377	Nonascul
872	15.8	83.2	180717	12	AC181132	AC181132 Strongylo	945	15.8	83.2	195833	12	AC168568	Strongylo
873	15.8	83.2	180889	12	AC079986	AC079986 Homo sapi	946	15.8	83.2	196422	12	AC147601	Homo sapi
874	15.8	83.2	181233	12	AC149832	AC149832 Zea mays	947	15.8	83.2	196518	5	AC116154	Homo sapi
875	15.8	83.2	181308	5	AL353148	AL353148 Human DNA	948	15.8	83.2	196567	6	AL928855	Mouse DNA
876	15.8	83.2	181847	6	AL512583	AL512583 Mouse DNA	949	15.8	83.2	196860	6	AC168279	Mus muscu
877	15.8	83.2	182038	5	AC073042	AC073042 Homo sapi	950	15.8	83.2	197031	12	AC019121	Homo sapi
878	15.8	83.2	182340	12	AC024417	AC024417 Homo sapi	951	15.8	83.2	197292	12	AC182010	Echinops
879	15.8	83.2	182431	12	AC151558	AC151558 Dasyapus n	952	15.8	83.2	197316	6	AC134899	Mus muscu
880	15.8	83.2	182520	6	AL627392	AL627392 Mouse DNA	953	15.8	83.2	197430	12	AC140273	Mus muscu
881	15.8	83.2	182599	12	AC022897	AC022897 Homo sapi	954	15.8	83.2	197565	11	CR450724	Zebrafish
882	15.8	83.2	182658	6	AC164549	AC164549 Mus muscu	955	15.8	83.2	197602	12	AC175711	Pan trogl
883	15.8	83.2	182912	5	AC068060	AC068060 Homo sapi	956	15.8	83.2	197646	6	AC136376	Mus muscu
884	15.8	83.2	182931	5	AC068856	AC068856 Homo sapi	957	15.8	83.2	197648	6	AC164400	Bos tauru
885	15.8	83.2	183054	6	AC154615	AC154615 Mus muscu	958	15.8	83.2	197693	12	AC146832	Bos tauru
886	15.8	83.2	183295	12	AC159805	AC159805 Glycine m	959	15.8	83.2	198006	6	AC119210	Mus muscu
887	15.8	83.2	183355	12	AC178035	AC178035 Strongylo	960	15.8	83.2	198051	12	AC152620	Bos tauru
888	15.8	83.2	183487	12	AC068828	AC068828 Homo sapi	961	15.8	83.2	198307	6	AC132432	Mus muscu
889	15.8	83.2	183577	12	AC011164	AC011164 Homo sapi	962	15.8	83.2	198318	12	CT486000	Mus muscu
890	15.8	83.2	183850	12	AC172889	AC172889 Phytophch	963	15.8	83.2	198336	5	AC010355	Homo sapi
891	15.8	83.2	184436	6	AL671870	AL671870 Mouse DNA	964	15.8	83.2	198440	12	CR925729	Danio rer
892	15.8	83.2	184700	6	AL671870	AL671870 Mouse DNA	965	15.8	83.2	198575	5	AC007495	Homo sapi
893	15.8	83.2	185091	12	AC015835	AC015835 Homo sapi	966	15.8	83.2	198635	12	AC121182	Rattus no
894	15.8	83.2	185121	12	AC079806	AC079806 Homo sapi	967	15.8	83.2	198693	12	AC166721	Bos tauru

```
c 968 15.8 83.2 198880 12 AC118266 Mus muscu
c 969 15.8 83.2 198885 12 CT033785 Mus muscu
c 970 15.8 83.2 199014 12 AC073849 Homo sapi
c 971 15.8 83.2 199191 6 AC158639 Mus muscu
c 972 15.8 83.2 199277 12 AC165893 Bos tauru
c 973 15.8 83.2 199384 5 AP003531 Homo sapi
c 974 15.8 83.2 199426 11 CR847506 Zebrafish
c 975 15.8 83.2 200098 12 AC091161 Homo sapi
c 976 15.8 83.2 200192 12 AC109885 Rattus no
c 977 15.8 83.2 200202 6 AC112154 Mus muscu
c 978 15.8 83.2 200321 11 BX510646 Zebrafish
c 979 15.8 83.2 200439 6 AC160032 Mus muscu
c 980 15.8 83.2 200591 6 CT025604 Mouse DNA
c 981 15.8 83.2 200647 12 AL954633 Mus muscu
c 982 15.8 83.2 200796 12 AC120910 Rattus no
c 983 15.8 83.2 200985 6 AC121787 Mus muscu
c 984 15.8 83.2 201133 12 AL807235 Homo sapi
c 985 15.8 83.2 201466 12 AC147573 Homo sapi
c 986 15.8 83.2 202162 12 AC129627 Rattus no
c 987 15.8 83.2 202342 6 AL671867 Mouse DNA
c 988 15.8 83.2 203403 6 AC157789 Mus muscu
c 989 15.8 83.2 203589 6 AC084383 Mus muscu
c 990 15.8 83.2 203626 12 AC138730 Pongo pyg
c 991 15.8 83.2 203732 12 AC129845 Rattus no
c 992 15.8 83.2 204078 12 AC178149 Strongylo
c 993 15.8 83.2 204462 12 AC155619 Zea mays
c 994 15.8 83.2 204498 6 AC159547 Mus muscu
c 995 15.8 83.2 205357 12 AC152018 Colobus g
c 996 15.8 83.2 205900 6 AC103610 Mus muscu
c 997 15.8 83.2 206004 4 AP008246 Oryza sat
c 998 15.8 83.2 206096 12 AC181811 Strongylo
c 999 15.8 83.2 206125 6 AC153382 Mus muscu
c1000 15.8 83.2 206246 12 AC132970 Rattus no
```

ALIGNMENTS

```
RESULT 1
BD233463/c 603 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION BD233463
VERSION BD233463.1 GI:33043233
KEYWORDS JP 2002519016-A/9.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 603)
Kato, S. and Kimura, T.
TITLE Human protein having hydrophobic domain and DNA encoding the same
JOURNAL Patent: JP 2002519016-A 9 02-JUL-2002;
SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
COMMENT OS Homo sapiens (human)
PN JP 2002519016-A/9
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO, TOMOKO KIMURA
PC C12N15/09, C07K14/47, C12N1/15, C12N1/19, C12N5/10, C12N5/00, C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
FH same
FT key Location/Qualifiers
1. .603
/organism="Homo sapiens (human)"
/locus="BD233463"
/db_xref="taxon:9606"
FEATURES
source
1. .603
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 603;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 591 CTTTACTTCATAGTCTTTG 573

RESULT 2
BD209699/c 1212 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Compositions isolated from skin cells and methods for their use.
ACCESSION BD209699
VERSION BD209699.1 GI:33019469
KEYWORDS JP 2002512798-A/171.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1212)
Strachan, L., Sleeman, M., Watson, J. D., Onrust, R., Kumble, A. and
Murison, J. G.
TITLE Compositions isolated from skin cells and methods for their use
JOURNAL Patent: JP 2002512798-A 171 08-MAY-2002;
GENESIS RESEARCH AND DEVELOPMENT CORP LTD
COMMENT OS Homo sapiens (human)
PN JP 2002512798-A/171
PD 08-MAY-2002
PF 29-APR-1999 JP 2000546009
PR 29-APR-1998 US 09/069726, 09-NOV-1998 US 09/188930 PI
LORNA STRACHAN, MATTHEW SLEEMAN, JAMES DOUGLAS WATSON, RENE PI
ONRUST,
PI ANAND KUMBLE, JAMES GREG MURISON
PC C12N15/09, A61K38/00, A61P9/00, A61P17/00, A61P29/00, A61P31/18, PC
A61P35/00,
PC
C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12N15/ PC
00, A61K37/02, C12N5/00
CC Compositions isolated from skin cells and methods for their
use.
FT key Location/Qualifiers
1. .1212
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
1. .1212
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 1212;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 1031 CTTTACTTCATAGTCTTTG 1013

RESULT 3
AR341505/c 1212 bp DNA linear PAT 17-AUG-2003
LOCUS
DEFINITION Sequence 249 from patent US 6573095.
ACCESSION AR341505
VERSION AR341505.1 GI:33733640
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
```

```

Unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Strachan,L.
TITLE Polynucleotides isolated from skin cells
JOURNAL Patent: US 6573095-A 249 03-JUN-2003;
        Genesis Research & Development Corporation Limited; Parnell;
        NZX;
FEATURES             Location/Qualifiers
   source             1..1212
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
Query Match          100.0%; Score 19; DB 2; Length 1212;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 1031 CTTTACTTCATAGTCTTTG 1013

RESULT 4
AX078375/c
LOCUS AX078375 1619 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 43 from Patent WO0107612.
ACCESSION AX078375
VERSION AX078375.1 GI:13158044
KEYWORDS
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
        Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
TITLE Receptors and associated proteins
JOURNAL Patent: WO 0107612-A 43 01-FEB-2001;
        Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
   source             1..1619
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 2681738CB1".
ORIGIN
Query Match          100.0%; Score 19; DB 2; Length 1619;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 1178 CTTTACTTCATAGTCTTTG 1160

RESULT 5
BD233473/c
LOCUS BD233473 1718 bp DNA linear PAT 17-JUL-2003
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION BD233473
VERSION BD233473.1 GI:33043243
KEYWORDS JP 2002519016-A/19.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1718)
AUTHORS Kato,S. and Kimura,T.
TITLE Human protein having hydrophobic domain and DNA encoding the same
JOURNAL Patent: JP 2002519016-A 19 02-JUL-2002;

```

```

SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
OS Homo sapiens (human)
PN JP 2002519016-A/19
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO,TOMOKO KIMURA
PC C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
CC same
FH Key 1..1718 Location/Qualifiers
FT source /organism="Homo sapiens (human)".
FEATURES             Location/Qualifiers
   source             1..1718
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
Query Match          100.0%; Score 19; DB 2; Length 1718;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 602 CTTTACTTCATAGTCTTTG 584

RESULT 6
AK172760/c
LOCUS AK172760 1826 bp mRNA linear PRI 07-MAY-2004
DEFINITION Homo sapiens cDNA FLJ3921 fis, clone COL02043, highly similar to
        Homo sapiens retinoic acid induced 3 (RAI3).
ACCESSION AK172760
VERSION AK172760.1 GI:47077732
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
        Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
        Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1826)
        Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2004) Sumio Sugano, Institute of Medical Science,
        University of Tokyo, Laboratory of Genome Structure, Human Genome
        Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
        (E-mail:flcdna@mail.ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
        Fax:81-3-5449-5416)
        NEDO human cDNA sequencing project supported by Ministry of
        Economy, Trade and Industry of Japan; cDNA full insert sequencing:
        Research Association for Biotechnology; cDNA library construction,
        5'- & 3'-end one pass sequencing; Department of Virology and Human
        Genome Center, Institute of Medical Science, University of Tokyo
        (partly supported by Science and Technology Agency).
FEATURES             Location/Qualifiers
   source             1..1826
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL02043"
                     /tissue_type="colon"
                     /clone_lib="COL"
                     /note="cloning vector pME18SFL3"
ORIGIN

```

```

Query Match      100.0%; Score 19; DB 5; Length 1826;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
Db 679 CTTTACTTCATAGTCCTTTG 661

RESULT 7
LOCUS CQ723177/c
DEFINITION Sequence 9111 from Patent WO02068579.
ACCESSION CQ723177
VERSION CQ723177.1 GI:42284034
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
  Kits, such as nucleic acid arrays, comprising a majority of
  humanexons or transcripts, for detecting expression and other uses
  thereof
  Patent: WO 02068579-A 9111 06-SEP-2002;
  PE Corporation (NY) (US)
FEATURES
    source
      1. .2290
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN

Query Match      100.0%; Score 19; DB 2; Length 2290;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
Db 1158 CTTTACTTCATAGTCCTTTG 1140

RESULT 8
LOCUS BC003665/c
DEFINITION Homo sapiens G protein-coupled receptor, family C, group 5, member
A, mRNA (cDNA clone MGC:923 IMAGE:2988011), complete cds.
ACCESSION BC003665
VERSION BC003665.2 GI:33872669
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2296)
  Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
  Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
  Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
  Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,
  Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
  Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
  Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
  Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
  Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
  McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
  Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
  Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
  Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
  Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

```

```

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Snailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2296)
Director MGC Project.
Direct Submission
Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:13177795.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahy, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 3 Row: 1 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 12056470.

FEATURES
    source
      1. .2296
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="MGC:923 IMAGE:2988011"
        /tissue_type="Colon, adenocarcinoma"
        /clone_lib="NIH MGC 15"
        /lab_host="DH10B-R"
        /notes="Vector: pOTB7"
      1. .2296
        /gene="GPCR5A"
        /note="synonyms: RAIG1, GPRC5A, RAI3"
        /db_xref="GeneID:9052"
        /db_xref="MIM:604138"
        95_..1168
        /gene="GPCR5A"
        /codon_start=1
        /product="G protein-coupled receptor, family C, group 5,
        member A"
        /protein_id="AAH03665.1"
        /db_xref="GI:13177796"
        /db_xref="GeneID:9052"
        /db_xref="MIM:604138"
        /translations="MATTVPDGGCRNGLSKSKYRLCDKAEAGIVLETATAGVVTVA
        FMTLTPILVCKQVDSNRRKMLPTQFLFGLVGLFGLTFAPIIGLDGSGTGPTRFLFG
        ILFSICSLAHAVSITKLVRGKPLSLVILGLAVGFSLVQDVIAIEVLVTNRT
        NVNFWELASPRNEDFVLLLTLYVLFMLTFLMSSFTFCGSGTGWKRGHAHLYTML
        LSAIWMATLMLLPDFRRWDITLSSALAANGVFLIAYVSPFWLLTKQNPM
        YPVEDAFCKPQLVKYSGVENRAYSQEITQGFETGDTLIYAPYSTHFLQNPQPOKE
        FSIPIRAHAMPSPYKDEVKKEGS"
ORIGIN

Query Match      100.0%; Score 19; DB 5; Length 2296;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
Db 1153 CTTTACTTCATAGTCCTTTG 1135

```

```

RESULT 9
AF506289/c
LOCUS Homo sapiens orphan G protein-coupling receptor PEIG-1 mRNA, PRI 15-JUL-2002
DEFINITION complete cds.
ACCESSION AF506289
VERSION AF506289.1 GI:21779962
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Terrett,J.A.
TITLE Diagnosis of carcinoma using raig1 polypeptides
JOURNAL Patent: WO 03087832-A 2 23-OCT-2003;
Oxford GlycoSciences (UK) Limited (GB)
FEATURES
source
1. .2302
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 19; DB 5; Length 2297;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
Db 1166 CTTTACTTCATAGTCTTTG 1148

RESULT 10
AX930411/c
LOCUS AX930411
DEFINITION Sequence 2 from Patent WO03087832.
ACCESSION AX930411
VERSION AX930411.1 GI:40312209

```

```

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Terrett,J.A.
TITLE Diagnosis of carcinoma using raig1 polypeptides
JOURNAL Patent: WO 03087832-A 2 23-OCT-2003;
Oxford GlycoSciences (UK) Limited (GB)
FEATURES
source
1. .2302
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 2302;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
Db 1158 CTTTACTTCATAGTCTTTG 1140

RESULT 11
AF095448/c
LOCUS AF095448
DEFINITION Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA,
complete cds.
ACCESSION AF095448
VERSION AF095448.1 GI:4063889
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Cheng,Y. and Lotan,R.
TITLE Molecular cloning and characterization of a novel retinoic
acid-inducible gene that encodes a putative G protein-coupled
receptor
JOURNAL J. Biol. Chem. 273 (52), 35008-35015 (1998)
PUBMED 9857033
REFERENCE
AUTHORS Cheng,Y. and Lotan,R.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1998) Tumor Biology, The University of Texas M.
D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX
77030, USA
FEATURES
source
1. .2302
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p12.3-p13"
1. .2302
/gene="RAIG1"
/notes="induced by retinoic acid"
100. .1173
/gene="RAIG1"
/codon_start=1
/product="putative G protein-coupled receptor"
/db_xref="GI:4063890"
/translation="MATTVPDGRNGLSKYRLCDKAEANGIVLETVATAGVVTVA
FMLTPLIVCKVQDSNRKMLPTQFLFLGLVGLFGLTFAFIIGLDGSTGPTFFLFG
ILFSICFSCLLAHAVSLTKLVGRKPLSLVILGLAVGFSLVQDVIATIEIVLTNNRT
LSIAIWMATITLMLPDRDWDPTILSSAANGWVFLLAYVSPFWLLTKQRNPMD
YPVEDAFCKVKKGYGVENRAYSQBEITQGETGTLYAPYSTHFLQNQPQKE
FSIPRAHWSPYKDYEVKKEGS"
gene
CDS
Qy 1 CTTTACTTCATAGTCTTTG 19
Db 1166 CTTTACTTCATAGTCTTTG 1148

RESULT 10
AX930411/c
LOCUS AX930411
DEFINITION Sequence 2 from Patent WO03087832.
ACCESSION AX930411
VERSION AX930411.1 GI:40312209

```

LSTAIWVWNTILMLDPDRRDDTTLSSALAANGVFLAYVSPFWLLTKORNPMW
YPVEDAFCKPQLVKKSYGVENRAYSQEEITQGFETGDTLYAPYSTHFLQNPQPOKE
FSIPRAHAWSPKYDEVKKEGS"

ORIGIN

Query Match 100.0%; Score 19; DB 5; Length 2302;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 1158 CTTTACTTCATAGCTTTG 1140

RESULT 12

BD156680/c
LOCUS BD156680 2446 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156680

VERSION BD156680.1 GI:27862438

KEYWORDS JP 2002191363-A/11523.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 2446)

Ota,T., Isoqai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 11523 09-JUL-2002;

HELIOS RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11523

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA,TAKAO ISOQAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU

PI SAITO,

PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,

PI KEIICHI NAGAI,TETSUJI OTSUKI

PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC

10, C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers (254). (1324).

FT CDS Location/Qualifiers

1. .2446

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 1312 CTTTACTTCATAGCTTTG 1294

RESULT 13

AX877483/c
LOCUS AX877483 2446 bp DNA linear PAT 17-DEC-2003

DEFINITION Sequence 12388 from Patent EP1074617.

ACCESSION AX877483

VERSION AX877483.1 GI:40032219

KEYWORDS Homo sapiens (human)

SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

1 Ota,T., Isoqai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primers for synthesizing full-length cDNA and their use

Patent: EP 1074617-A 12388 07-FEB-2001;

Research Association for Biotechnology (JP)

Location/Qualifiers

source

1. .2446

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

254. .1327

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE89746.1"

/db_xref="GI:40032220"

/translation="MATTFDGCGRNGKSKYYFLCDKAEAWGIVLETVATAGVVTVA

FMLTLFLVKQVDSNRRKMLPQFLGLVGLGIFGLTAFIILGDSGTPTFFLFG

ILFSICFCLLAHVSATKLVGRKPLSLVILGLAVGFSLVQDVIAIEIVLTMNRT

NNVFSLSAPRRNEDFVLLITVFLMALTFMLMSFTFCGSGTKWRGHAIYLTML

LSTAIWVWNTILMLDPDRRDDTTLSSALAANGVFLAYVSPFWLLTKORNPMW

YPVEDAFCKPQLVKKSYGVENRAYSQEEITQGFETGDTLYAPYSTHFLQNPQPOKE

FSIPRAHAWSPKYDEVKKEGS"

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 2446;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 1312 CTTTACTTCATAGCTTTG 1294

RESULT 14

AK001761/c

LOCUS AK001761 2446 bp mRNA linear PRI 20-JAN-2006

DEFINITION Homo sapiens cDNA FLJ10899 fis, clone NT2RP5003506.

ACCESSION AK001761

VERSION AK001761.1 GI:7023229

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

1 Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,

Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,

Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,

Nagahari,K., Murakami,K., Yaeuda,T., Iwayanagi,T., Wagatsuma,M.,

Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,

Sugawara,M., Takahashi,M., Kanda,K., Kamiara,K., Katsuta,N., Sato,K.,

Kikkawa,B., Omura,Y., Abe,K., Kimihara,K., Ishibashi,T., Yamashita,H.,

Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,

Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,

Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,

Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,

Hara,H., Tanase,T.O., Nomura,Y., Kanehori,K., Komai,F., Hara,R.,

Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,

Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihiro,T.,

Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,

Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,

Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,

Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,

Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,

Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,

Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,

Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,

Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,

Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

2 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Negai, K., Sugano, S., Ishibashi, T., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Kanehori, K.
NEDO human cDNA sequencing project
Unpublished

3 (bases 1 to 2446)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana Kiearazu Chiba 292-0812, Japan (E-mail: flj-cdn@nifty.com, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
Location/Qualifiers
1. .2446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="NT2RP5003506"
/cell_lines="NT2"
/cell_type="teratocarcinoma"
/clone_libs="NT2RP5"
/note="Cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
254. .1327
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91890.1"
/db_xref="GI:7023230"
/translation="MATVTPDGRNGLSKYYRLCDKAEAWGIVLETVATAGVVTSA
FMLTPIIVCKVQDSNRKMLPTQFLFLVGLTFLGTFAPIIGLDGSGTGRFLFG
ILFSICFSLAAVSLTKLVGRKPLSLVLGLAVGFSLVQDVIAIEYVLTMNRT
NNVNSFELSAPRNEDFVLLITVYVFLMALTFELMSFTFCGSGTGWKRGHAIYLTWL
LSIAIWAIVNTLLMLPDRDRDDTLSSALAANGVFWLLAYVSFEFWLLTYQRNPD
YVDEDAFCPOLVKVSKYGVENRAYSQBEITGFTGDTLYAPYSTHFLQNPQPKQE
FSIPRAHWSPYKDYEVKKEGS"

CDS
Query Match 100.0%; Score 19; DB 5; Length 2446;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCTTTG 1294

RESULT 15
LOCUS CQ981495 2456 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 350 from Patent EP1498424.
ACCESSION CQ981495
VERSION CQ981495.1 GI:58190785
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 Rosenthal, A., Hermann, K., Heiden, E., Pilarczyk, C., Bruemendorf, T., Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X. Human nucleic acid sequences from lung tumours
Patent: EP 1498424-A 350 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE); Rosenthal, Andre (DE)
Location/Qualifiers
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCTTTG 1294

RESULT 16
LOCUS DD210040/c 2456 bp DNA linear PAT 19-JAN-2006
DEFINITION Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer.
ACCESSION DD210040
VERSION DD210040.1 GI:85654022
KEYWORDS JP 2005518782-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2456)
Zlotnik, A., Mack, D.H., Agiz, N., Gish, K.C., Hebeji, P.A., Wilson, K.E. and Afar, D.
Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer
Patent: JP 2005518782-A 34 30-JUN-2005;
PROTEIN DESIGN LABS INC
OS Homo sapiens
PN JP 2005518782-A/34
PD 30-JUN-2005
PF 17-SEP-2002 JP 2003529912
PR 12-APR-2002 US 60/372246, 08-FEB-2002 US 60/355257, PR 08-FEB-2002 US 60/355145, 13-NOV-2001 US 60/350666, PR 08-FEB-2001 US 60/323887, 17-SEP-2001 US 60/323469 PI albert
zlotnik, david h mack, natasha agiz, kurt c gish, peter a pi hebeji, pi keith e wilson, daniel afar
CC
FH Key Location/Qualifiers.
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCTTTG 1294


```

RESULT 17
AX549168/c
LOCUS
DEFINITION Sequence 453 from Patent WO02061087.
ACCESSION AX549168
VERSION AX549168.1 GI:25813894
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1
AUTHORS
TITLE Antigenic peptides, such as for G protein-coupled receptors
(GPCRs) antibodies thereto, and systems for identifying such
antigenic peptides
JOURNAL Patent: WO 02061087-A 453 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 2456;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 18
AK122672/c
LOCUS
DEFINITION Homo sapiens cDNA FLJ16117 fis, clone ASTRO2003632, highly similar
to Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA.
ACCESSION AK122672
VERSION AK122672.1 GI:34527861
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1
AUTHORS
TITLE Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Murakawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Haraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T.O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Saeaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terasima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, K., Kawakami, T.,

```

```

Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
JOURNAL
PUBMED 14702039
REFERENCE
2
AUTHORS Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M.,
Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 3057)
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="ASTRO2003632"
/cell_type="normal astrocytes (NEA5732)"
/clone_lib="ASTRO2"
/notes="cloning vector: pME18SFL3
primary culture, normal astrocytes"
FEATURES
source
1. .3057
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="ASTRO2003632"
/cell_type="normal astrocytes (NEA5732)"
/clone_lib="ASTRO2"
/notes="cloning vector: pME18SFL3
primary culture, normal astrocytes"
ORIGIN
Query Match 100.0%; Score 19; DB 5; Length 3057;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1949 CTTTACTTCATAGTCCTTTG 1931

RESULT 19
AX188348/c
LOCUS
DEFINITION Sequence 4043 from Patent WO0142467.
ACCESSION AX188348
VERSION AX188348.1 GI:15139821
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1
AUTHORS
TITLE Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer

```

JOURNAL Patent: WO 0142457-A 4043 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
source 1..3371
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 3371;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1225 CTTTACTTCATAGTCTTTG 1207
RESULT 20
CQ894732/C
LOCUS CQ894732 6730 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 42 from Patent EP1471075.
ACCESSION CQ894732
VERSION CQ894732.1 GI:55467481
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Rosenthal, A.D., Pilarsky, C., Dahl, E., Specht, T., Brueemendorf, T.,
Lichtner, R., Staub, E., Roepcke, S. and Li, X.I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: EP 1471075-A 42 27-OCT-2004;
Hinzmann, Bernd (DE); Rosenthal, Andre (DE); Pilarsky, Christian
(DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie
(DE)
FEATURES Location/Qualifiers
source 1..6730
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 6730;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCTTTG 1294
RESULT 21
AC007688
LOCUS AC007688 161577 bp DNA linear PRI 30-AUG-2002
DEFINITION Homo sapiens 12 BAC RPC111-392P7 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
ACCESSION AC007688
VERSION AC007688.15 GI:5815499
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 161577)
AUTHORS Muzny, D., Arenson, A.D., Bouck, J., Bunac, C., Chen, Z., Ding, Y.,
Dugan, S., Durbin, J., Forcum, J., Garcia, C., Gorrell, J.H.,
Gorrell, L.L., Hernandez, J., Issat, A., Jackson, L., Kneitz, S.,
Kondejewski, N., Lau, S., Leal, B., Lee, E., Lichtarge, O., Liu, W.,
Logan, O., Lu, J., Marondel, I., Martinez, C., Merscher, S., Miller, A.,

Montgomery, K., Oswal, G., Pampell, L.R., Parish, B.J., Perez, L.,
Rashid, N.D., Rives, C., Scherer, S.E., Shen, H., Shim, C., Simon, M.,
Vo, Q., Williamson, A., Worley, K.C., Xiang, A.M., Yang, R., Yu, W.,
Zhou, X., Kuchera, R., Nelson, D. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 1, 1999 this sequence version replaced gi:5757565.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 161577
Phrap values in estimate: 160751

Average error rate (BCM-Phrap estimate): 0.000163681
 Fraction of Phrap values less than 40 : 0.0376047
 Number of consensus changing edits: 30
 Number of N's in consensus : 0

```

----- Consensus changing edits -----
Position      Original+Context      Edited+Context
7033          acctgcgcgt (n) cgcgcgcctt
47567         aaaaaaaa (n) gaaaaataat
51135         aaaaaaaa (n) gaaaaataat
51135         aagaagaa (n) aagaagaa
75582         aaaaaaaa (n) aanggatgtt
75585         aaaaaaaa (n) gaaggttcgc
75667         ctaagcaga (n) taagatctta
75752         tttaaatag (g) gctttgctt
84017         ggaaggaag (n) aagaagagag
85227         tttttttt (n) tgtttttttt
99681         ggaagtcag (n) atgcagtgag
111307        aatctctat (n) ccgaattca
111374        actagtatac (n) atctttttt
111378        gtatacnatc (n) tttttttt
112621        aaaaaaaa (n) ccatcctaga
135812        attcaccctc (c) tttttttt
137207        ttgcgggac (n) cgcaccacg
137218        cgcaccacg (n) ctggctaagt
145113        gaaagtga (n) natgtagaat
145114        caagtgaan (n) atgtagaata
145232        ggcaccgtg (n) ntacnccag
145233        gacccgtgn (n) tcacnccag
145238        ggtgntcac (n) ccagtaatcc
145321        gcaacatggt (t) nmccccatc
145322        caacatggtt (n) nmccccatc
145323        aacatggttn (n) nmccccatc
145324        atcaggttnn (n) cccatctct
145363        gtccaccag (n) gtggtggcgt
145377        gggcggtga (t) ttgacgctt
145397        tgcctacttg (n) gggggggag
145505        aaaaaaaaa (n) gaaaaaaaa

```

----- Distribution of Quality < 40 Bases -----

#	Phrap Value Range									
	5	10	15	20	25	30	35	40		
1000	*	*	*	*	*	*	*	*		
900	*	*	*	*	*	*	*	*		
800	*	*	*	*	*	*	*	*		
700	*	*	*	*	*	*	*	*		
600	*	*	*	*	*	*	*	*		
500	*	*	*	*	*	*	*	*		
400	*	*	*	*	*	*	*	*		
300	*	*	*	*	*	*	*	*		
200	*	*	*	*	*	*	*	*		
100	*	*	*	*	*	*	*	*		
0	*	*	*	*	*	*	*	*		

Version: 1.01 qxf0.
 FEATURES
 source
 Location/Qualifiers
 1. 161577
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /clone="RP11-392P7"
 complement(3. .65)
 /rpt_family="MER4B"
 161. .471
 /rpt_family="AluSx"
 complement(607. .702)
 /rpt_family="MIR"

```

repeat_region 703. .838
                /rpt_family="FLAM_C"
repeat_region complement(839. .916)
                /rpt_family="MIR"
repeat_region complement(1205. .1377)
                /rpt_family="MER104"
STS
1388. .1525
  /standard_name="WIAP-759-STS"
  /db_xref="dbSTS:65358"
  1683. .1971
    /rpt_family="AluJo"
repeat_region 1974. .2261
    /rpt_family="AluJb"
repeat_region 2486. .2593
    /rpt_family="MIR"
misc_feature 2610. .3214
  /note="Region: Unigene cluster similar to AA056332 and
  AI074576"
STS
2712. .2823
  /standard_name="SHGC-44583"
  /db_xref="dbSTS:48451"
  3185. .3491
    /rpt_family="AluSq"
repeat_region complement(6397. .6537)
    /rpt_family="L2"
repeat_region 6967. .7158
    /rpt_family="CCCCG)n"
repeat_region 7094. .7240
    /rpt_family="CCCCG)n"
repeat_region complement(7770. .8250)
    /rpt_family="MER44B"
repeat_region 8831. .9073
    /rpt_family="MER102"
repeat_region complement(9075. .9303)
    /rpt_family="L2"
repeat_region 9671. .10135
    /rpt_family="LTR33A"

```

Query Match 100.0%; Score 19; DB 5; Length 161577;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 94870 CTTTACTTCATAGTCCTTTG 94888

RESULT 22
 RP43035C13 157305 bp DNA linear PRI 03-FEB-2005
 LOCUS Pan troglodytes chromosome 22 BAC RP43-035C13, complete sequence.
 DEFINITION
 ACCESSION AL954230
 VERSION AL954230.2 GI:38453651
 KEYWORDS HTG.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Pan.

REFERENCE
 1 Watanabe,H., Fujiyama,A., Hattori,M., Taylor,T.D., Toyoda,A.,
 Kuroki,Y., Noguchi,H., Benkhalil,A., Lehrach,H., Sudbrak,R.,
 Kube,M., Taenzer,S., Galocz,P., Platzer,M., Scharfe,M.,
 Nordliek,G., Blocker,H., Hellmann,I., Khaitovich,P., Paabo,S.,
 Reinhardt,R., Zheng,H.J., Zhang,X.L., Zhu,G.F., Wang,B.F., Fu,G.,
 Ren,S.X., Zhao,G.P., Chen,Z., Lee,Y.S., Cheong,J.E., Choi,S.H.,
 Wu,K.M., Liu,T.T., Hsiao,K.J., Tsai,S.F., Kim,C.G., Oota,S.,
 Kitano,T., Kohara,Y., Saitou,N., Park,H.S., Wang,S.Y., Taspo,M.L.
 and Sakaki,Y.
 The Chimpanzee Chromosome 22 Sequencing Consortium
 DNA sequence and comparative analysis of chimpanzee chromosome 22
 Nature 429 (6990), 382-388 (2004)
 PUBLISHED 15164055


```
exon      complement(7529..7853)
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 86.000%, comment = excellent"
complement(7546..7853)
/inference="non-experimental evidence, no additional
details recorded"
/note="Xpound exon prediction, score = 98% (0%)"
complement(7563..7853)
/inference="non-experimental evidence, no additional
details recorded"
/note="MZEF, score = 72.2%"
complement(8472..8544)
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 49.000%, comment = marginal"
complement(8772..8977)
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 60.000%, comment = good"
complement(8780..8977)
/inference="non-experimental evidence, no additional
details recorded"
/note="Genscan, score = 1.30%, comment = Internal_exon 198
bp frame: 2 phase: 0"
complement(9028..9216)
/inference="non-experimental evidence, no additional
details recorded"
/note="Genscan, score = 13.54%, comment = Internal_exon
189 bp frame: 1 phase: 0"
complement(9028..9216)
/inference="non-experimental evidence, no additional
details recorded"
/note="MZEF, score = 96.1%"
complement(9028..9216)
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 92.000%, comment = excellent"
complement(9842..9963)
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="MIR"
11055..11354
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="AluY"
11606..11795
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="MER63A"
12023..12324
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="AluX"
complement(12784..12882)
/inference="non-experimental evidence, no additional
details recorded"
/note="Genscan, score = 0.56%, comment = Internal_exon 99
bp frame: 1 phase: 0"
complement(12784..12882)
/inference="non-experimental evidence, no additional
details recorded"
/note="MZEF, score = 96.1%"
12814..12872
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 59.000%, comment = good shadow"
complement(13508..13527)
/inference="non-experimental evidence, no additional
details recorded"
/note="Xpound exon prediction, score = 67% (0%)"
complement(14478..14645)
/inference="non-experimental evidence, no additional
```

```
exon      details recorded"
/note="MZEF, score = 86.2%"
complement(14478..14641)
/inference="non-experimental evidence, no additional
details recorded"
/note="Genscan, score = 9.80%, comment = Internal_exon 164
bp frame: 0 phase: 2"
complement(14478..14641)
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 94.000%, comment = excellent"
complement(14478..14590)
/inference="non-experimental evidence, no additional
details recorded"
/note="Xpound exon prediction, score = 86% (0%)"
complement(16046..16163)
/inference="non-experimental evidence, no additional
details recorded"
/note="Genscan, score = 6.30%, comment = Internal_exon 118
bp frame: 0 phase: 1"
16257..16330
/inference="non-experimental evidence, no additional
details recorded"
/note="GRAIL, score = 43.000%, comment = marginal shadow"
18126..18259
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="MER5A"
19077..19876
/inference="non-experimental evidence, no additional
details recorded"
19945..20070
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="L1MC3"
20072..20405
/inference="non-experimental evidence, no additional
details recorded"
/note="MZEF, score = 84.2%"
20106..20254
/inference="non-experimental evidence, no additional
details recorded"
complement(20145..20172)
/inference="non-experimental evidence, no additional
details recorded"
/note="Xpound exon prediction, score = 73% (0%)"
complement(20491..20855)
/inference="non-experimental evidence, no additional
details recorded"
/rpt_family="THE1B"
20859..20899
repeat_region

Query Match      94.7%; Score 18; DB 5; Length 159424;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTT 18
        |||||
Db      76655 CTTTACTTCATAGTCTTT 76672

RESULT 24
AC087481      179400 bp      DNA      linear      PRI 25-JUN-2002
LOCUS      AC087481
DEFINITION      Homo sapiens chromosome 15, clone RP11-540B6, complete sequence.
ACCESSION      AC087481
VERSION      AC087481.9
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 179400)

AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens chromosome 15, clone RP11-540B6

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 179400)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schupback, R., Schauer, B., Stange-Thomann, N., Stojanovic, N., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 179400)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Mlenga, V., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, B., Schupback, R., Seaman, S., Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (14-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

4 (bases 1 to 179400)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Mlenga, V., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N.,

Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, B., Schupback, R., Seaman, S., Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 25, 2002 this sequence version replaced gi:21426184.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11939

Center clone name: 540_B_6

Only the first 179.4 kilobases of this clone are being submitted.

The remainder overlaps accession number AC004460 [WashU clone].

FEATURES

source

1. 179400
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clones="RP11-540B6"
/clone_lib="RPC1-11 Human Male BAC"
complement (1.. 459)
/rpt_family="L1MC4a"
141..145
/note="<30 qual SNGL region"
complement (500.. 646)
/rpt_family="L1MC3"
complement (917.. 1123)
/rpt_family="MERS8A"
complement (1403.. 1687)
/rpt_family="AluSx"
1808.. 1820
/note="<30 qual SNGL region"
1830.. 1834
/note="<30 qual SNGL region"
1852.. 1872
/note="<30 qual SNGL region"
2190.. 2356
/rpt_family="MIR"
complement (2576.. 2758)
/rpt_family="L2"
3842.. 3911
/rpt_family="L2"
complement (4527.. 4846)
/rpt_family="AluJo"
7737.. 7796
/rpt_family="L3"
complement (8202.. 8334)
/rpt_family="L2"
10511.. 10809
/rpt_family="L2"
complement (12144.. 12355)
/rpt_family="MIR"
12395.. 12437
/rpt_family="(TG)n"
complement (12712.. 12851)
/rpt_family="L2"
13689.. 13717
/rpt_family="AT_rich"

repeat_region

unsure

repeat_region

repeat_region

repeat_region

unsure

unsure

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```

repeat_region 15823..15909
/rpt_family="MER5B"
repeat_region Complement(16744..17080)
/rpt_family="L1M4"
repeat_region 17280..17331
/rpt_family="(TCTA)n"
repeat_region Complement(17332..17622)
/rpt_family="AluI"
repeat_region 17767..17894
/rpt_family="MER3"
repeat_region 19432..19731
/rpt_family="MER3"
repeat_region 19824..19953
/rpt_family="AluSx"
repeat_region 20195..20238
/rpt_family="(GGA)n"
repeat_region 20195..20238
/rpt_family="AT rich"
repeat_region Complement(20286..20650)
/rpt_family="L1MCS"
repeat_region Complement(20816..21056)
/rpt_family="L1MCS"
repeat_region 23098..23253
/rpt_family="L2"
repeat_region Complement(23254..23531)
/rpt_family="AluSx"
repeat_region 23532..23809
/rpt_family="L2"
repeat_region 24283..24529
/rpt_family="L2"
repeat_region 24621..24723
/rpt_family="U6"
repeat_region 25405..25702
/rpt_family="AluSg"
repeat_region 25946..26049
/rpt_family="L2"
repeat_region Complement(27440..27819)
/rpt_family="L1MD2"
repeat_region 27883..28175
/rpt_family="AluSx"
repeat_region 28188..28303
/rpt_family="FLAM_C"
repeat_region 28314..28606
/rpt_family="AluJo"
repeat_region 28652..28759
/rpt_family="Triggers"
repeat_region 29174..29468
/rpt_family="AluSx"

Query Match 94.7% Score 18; DB 5; Length 179400;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTT 18
|||||
Db 115416 CTTTACTTCATAGTCCTT 115433

RESULT 25
PTB144012
LOCUS PTB144012 180638 bp DNA linear PRI 03-FEB-2005
DEFINITION Pan troglodytes chromosome 22 BAC PTB-144012, complete sequence.
ACCESSION AL954231
VERSION AL954231.3 GI:38153749
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
REFERENCE 1
Watanabe,H., Fujiyama,A., Hattori,M., Taylor,T.D., Toyoda,A.,
Kuroki,Y., Noguchi,H., Benkhalil,A., Leirach,H., Sudbrak,R.,
Kube,M., Taenzler,S., Galoczky,P., Platzner,M., Scharfe,M.,
Nordsiek,G., Blocker,H., Hellmann,I., Khaitovich,P., Paabo,S.,

```

```

Reinhardt, R., Zheng, H.J., Zhang, X.L., Zhu, G.F., Wang, B.F., Fu, G.,
Ren, S.X., Zhao, G.P., Chen, Z., Lee, Y.S., Cheong, J.E., Choi, S.H.,
Wu, K.M., Liu, T.T., Hsiao, K.J., Tsai, S.F., Kim, C.G., Oota, S.,
Kitano, T., Kohara, Y., Saitou, N., Park, H.S., Wang, S.Y., Yaspo, M.L.
and Sakaki, Y.
The Chimpanzee Chromosome 22 Sequencing Consortium
DNA sequence and comparative analysis of chimpanzee chromosome 22
Nature 429 (6990), 382-388 (2004)
15164055
2 (bases 1 to 180638)
The Chimpanzee Chromosome 22 Sequencing Consortium.
Chimpanzee chromosome 22 genomic sequence
Unpublished
3 (bases 1 to 180638)
Schafie, M., Berg, C., Conrad, A., Hornischer, K., Loehnert, T.H.,
Ludwig, M., Thies, S., Weber, K. and Bloesker, H.
Direct Submission
Submitted (25-JUN-2003) GBF, Dept. of Genome Analysis, Mascheroder
Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
On Nov 3, 2003 this sequence version replaced gi:37620132.
The Chimpanzee Chromosome 22 Sequencing Consortium consists of :
*Chinese National Human Genome Center at Shanghai,
Shanghai, China
*GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRIIBB Genome Research
Center, Daejeon, Korea;
*Max-Planck Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
*****
+ Analysis and annotation were performed with the automatic
+ 'first-pass' annotation and submission tool
+ 'AnnoMitter' (Hornischer & Bloesker).
+ Programs used by 'AnnoMitter':
*****
----- Genome Center
Center: GBF, Braunschweig
Center code: GBF
Web site: http://genome.gbf.de/
Contact: info.genome@gbf.de
----- Project Information
Center project name:
Center clone name: PTB-144012
----- Summary Statistics
Sequencing vector: ##;
Chemistry: Dye-terminator-amersham: ## of reads
Chemistry: Dye-primer-amersham: ## of reads
Assembly program: ##
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: ##; agarose-fp estimation
Estimated insert size: 180638; sum-of-contigs estimation
*****
*
FEATURES
source Location/Qualifiers
1..180638
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-144012"
1..180638
/misc_feature /note="assembly_fragment-clone_end:T7-vector_side:left
assembly_fragment-clone_end:SP6-vector_side:right"
ORIGIN
Query Match 94.7% Score 18; DB 5; Length 180638;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


[illegible]

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE JOURNAL REFERENCE

2 (bases 1 to 206082)

Worley, K. C.

Direct Submission

TITLE JOURNAL

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 206082)

Rat Genome Sequencing Consortium.

Direct Submission

TITLE JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23264229.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College Of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GGPI

Center clone name: CH230-183A24

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 193499 bases at least Q40

Consensus quality: 195436 bases at least Q30

Consensus quality: 196841 bases at least Q20

Estimated insert size: 201950; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 204625: contig of 204625 bp in length

* 204626 204725: gap of unknown length

* 204726 206082: contig of 1357 bp in length.

FEATURES

source

Location/Qualifiers

1..206082

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-183A24"

misc_feature

1..1533

/note="wgs_end_extension"

clone_end:T7"

2990..4234

misc_feature

/note="wgs_end_extension"

clone_end:T7"

5188..5661

misc_feature

/note="clone_boundary
clone_end:T7
site:ECORI
end_sequence:BH339247"
complement(203280..203920)
/note="clone_boundary
clone_end:Sp6
site:ECORI
end_sequence:BH339260"
204626..204725
/estimated_length=unknown

misc_feature

204626..204725

Query Match 94.7%; Score 18; DB 12; Length 206082;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19

||||||||||||||||

Db 172840 TTTACTTCATAGTCTTTG 172823

||||||||||||||||

RESULT 29

CT027583/c

LOCUS

DEFINITION

Danio rerio chromosome 2 clone DKEY-23H1, WORKING DRAFT SEQUENCE, 7

unordered pieces.

ACCSSION

CT027583

VERSION

CT027583.4 GI:76061021

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

SOURCE

Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 249140)

McLaren, S.

Direct Submission

Submitted (02-FEB-2006) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfish-help@sanger.ac.uk Clone requests:

http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

----- Project Information

Center project name: zK23H1

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 246887 bases at least Q40

Consensus quality: 247347 bases at least Q30

Consensus quality: 247722 bases at least Q20

Insert size: 248540; sum-of-contigs

Insert size: 250444; 3.5% error; agarose-fp

Quality coverage: 7.80x in Q20 bases; sum-of-contigs Quality

coverage: 7.78x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently

* consists of 7 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved

* 1 63859: contig of 63859 bp in length

* 63860 63959: gap of 100 bp

* 63960 68546: contig of 4587 bp in length

* 68547 68646: gap of 100 bp

* 68647 100429: contig of 31783 bp in length

```

* 100430 100529: gap of 100 bp
* 100530 107413: contig of 6884 bp in length
* 107414 107513: gap of 100 bp
* 107514 191566: contig of 84053 bp in length
* 191567 191666: gap of 100 bp
* 191667 224283: contig of 32616 bp in length
* 224283 224383: gap of 100 bp
* 224383 249140: contig of 24758 bp in length.

FEATURES
    source
        1. .249140
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="RZPD:HUKB735H0123Q"
            /db_xref="taxon:7955"
            /chromosome="2"
            /clone="DKEY-23H1"
            /clone_lib="DanioKey"
        1. .63859
            /note="assembly_fragment:01360
            fragment_chain:1"
            63960. .68546
            /note="assembly_fragment:00017
            fragment_chain:1"
            68647. .100429
            /note="assembly_fragment:00932
            fragment_chain:1"
            100530. .107413
            /note="assembly_fragment:00080
            fragment_chain:1"
            107514. .191566
            /note="assembly_fragment:02187
            fragment_chain:2"
            191667. .224282
            /note="assembly_fragment:00549
            fragment_chain:2"
            224383. .249140
            /note="assembly_fragment:00173.0"

ORIGIN
    Query Match          94.7%; Score 18; DB 12; Length 249140;
    Best Local Similarity 100.0%; Pred. No. 1.7e+02;
    Matches 18, Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTT 18
    |||||
Db 176396 CTTTACTTCATAGTCCTTT 176379

RESULT 30
HS21C081
LOCUS HS21C081 340000 bp DNA linear PRI 16-APR-2005
DEFINITION Homo sapiens chromosome 21 segment HS21C081.
ACCESSION AL163281 AP001736 BA000005
VERSION AL163281.2 GI:7717373
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 340000)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schilhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Aakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuoyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstieck,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichert,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and

```

```

Yaspo,M.L.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e.mail: sakaki@gsc.riken.go.jp
* URL: http://hgp.gsc.riken.go.jp/
and
* Institute of Molecular Biotechnology, Genome Analysis, *
Beutenbergstrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e.mail: shimizu@mb-med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/
and
* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
* Ihnestrasse 73, D-14195 Berlin, Germany,
* e.mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.
Location/Qualifiers
1. .340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
<1. .1000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
<1. .147816
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC-291B3, 5' partial"
/clone_lib="BAC library"
/note="Accession No. AF064857"
134734. .167600
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="O7G4"
/clone_lib="LL21NCO2-Q Cosmid library"
/note="Accession No. AF045449"
154163. .300059
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"

```

```

/clones="P31P10"
/clone_lib="RPC11,3-5 PAC library"
/notes="Accession No. AF064862"
235367. .340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P58D10, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/notes="Accession No. AF064865"
2494. .2520
/notes="(CAAAA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2620. .3399
/notes="L1MC4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(5949. .6158)
/notes="AluJb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6294. .6451)
/notes="L1MB5"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
7710. .7717
/notes="Not1 site"
7942. .8206
/notes="AluJc"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8367. .8439)
/notes="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
8568. .8882
/notes="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11216. .11244)
/notes="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
complement(11245. .11259)
/notes="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
11262. .11331
/notes="MER16"
/rpt_family="DNA/Mariner"
/rpt_type=DISPERSED
complement(11341. .11465)
/notes="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
13420. .13527
/notes="(TG)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
13716. .14123
/notes="WL1D"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
16010. .16082
/notes="MER67A"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
16083. .16589
/notes="MER67B"
/rpt_family="LTR/MER4-group"

```

```

repeat_region
16623. .16789
/notes="MER67A"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
17023. .17083
/notes="AluSg"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(17728. .18041)
/notes="AluJb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(19405. .19488)
/notes="Alu"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(19506. .20061)
/notes="L1MEC"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(20236. .20397)
/notes="HERVL"
/rpt_family="LTR/Retroviral"
/rpt_type=DISPERSED
complement(20500. .20848)
/notes="HERVL"
/rpt_family="LTR/Retroviral"
/rpt_type=DISPERSED
complement(20944. .21212)
/notes="HERV16"

Query Match          94.7%; Score 18; DB 5; Length 340000;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTT 18
    |||||
Db 312022 CTTTACTTCATAGTCTTT 312039

RESULT 31
DQ271038/c
LOCUS      DQ271038          762 bp    DNA    linear    MAM 19-JAN-2006
DEFINITION Mustela vison clone VIIISN11 microsatellite sequence.
ACCESSION  DQ271038
VERSION    DQ271038.1  GI:82797063
KEYWORDS
SOURCE     Mustela vison (American mink)
ORGANISM  Mustela vison
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia;
            Mustelidae; Mustelinae; Mustela.
REFERENCE  1 (bases 1 to 762)
AUTHORS   Anistoraei,R. and Christensen,K.
TITLE     Isolation and characterization of 79 microsatellite makers for the
JOURNAL   American mink (Mustela vison)
REFERENCE  2 (bases 1 to 762)
AUTHORS   Anistoraei,R. and Christensen,K.
TITLE     Direct Submission
JOURNAL   Submitted (31-OCT-2005) Division of Animal Genetics, Department of
            Animal and Veterinary Basic Sciences, The Royal Veterinary and
            Agricultural University, Groennegaardsvej 3, Frederiksberg C 1870,
            Denmark
FEATURES
            source
            1. .762
            /organism="Mustela vison"
            /mol_type="genomic DNA"
            /db_xref="taxon:9667"
            /clone="VIIISN11"
            1. .762
            /note="microsatellite"

repeat_region

```

ORIGIN		/rpt_type=tandem	
Query Match Best Local Similarity Matches	91.6%;	Score 17.4;	DB 14; Length 762;
	94.7%;	Pred. No. 2.4e+02;	
	18; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	1	CTTTACTTCATAGCTTTTG	19
Db	600	CTTTACTTCATAGTTTTTG	582
RESULT 32			
SCYGR131W		1744 bp	DNA linear PLN 11-AUG-1997
LOCUS			
DEFINITION			S.cerevisiae chromosome VII reading frame ORF YGR131w.
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
			Saccharomyces cerevisiae (baker's yeast)
			Saccharomyces cerevisiae
			Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
			Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE			
AUTHORS			van Dyck,L., Skala,J., de Wergifosse,P., Purnelle,B., Talla,E.,
			Nawrocki,A., Del Bino,S. and Goffeau,A.
JOURNAL			Unpublished
REFERENCE			
AUTHORS			2 (bases 1 to 1744)
			MIPS.
TITLE			Direct Submission
JOURNAL			Submitted (14-MAY-1996) Data collected by MIPS on behalf of the
			European yeast chromosome VII sequencing project. MIPS at the
			Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
			Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
FEATURES			
source			
			Location/Qualifiers
			1..1744
			/organism="Saccharomyces cerevisiae"
			/mol_type="genomic DNA"
			/db_xref="taxon:4932"
			/chromosome="VII"
			complement(744..754)
			/note="ARS-consensus"
			882..1406
			/note="unnamed protein product; ORF YGR131w"
			/codon_start=1
			/protein_id="CAA97144.1"
			/db_xref="GI:1323217"
			/db_xref="GOA:P53279"
			/db_xref="SGD:S0003363"
			/db_xref="UniProtKB/Swiss-Prot:P53279"
			/translation="MISAADNLVRIINAVFLIISIGLSGLIGTQTQTKHSRVNFCMFA
			AVGLVTDLSLYGLANFWISLTYPAILLVLDPLNFIFTFVAATLAVGIRCHSKNKT
			YLSQNKIQSSSRCHSQSAQVAFFYFCFLFIKVATMGWMQNGFGSGNTGFSRR
			RARRQMGITISQV"
misc_feature			
			complement(744..754)
			/note="ARS-consensus"
			882..1406
CDS			
			/note="unnamed protein product; ORF YGR131w"
			/codon_start=1
			/protein_id="CAA97144.1"
			/db_xref="GI:1323217"
			/db_xref="GOA:P53279"
			/db_xref="SGD:S0003363"
			/db_xref="UniProtKB/Swiss-Prot:P53279"
			/translation="MISAADNLVRIINAVFLIISIGLSGLIGTQTQTKHSRVNFCMFA
			AVGLVTDLSLYGLANFWISLTYPAILLVLDPLNFIFTFVAATLAVGIRCHSKNKT
			YLSQNKIQSSSRCHSQSAQVAFFYFCFLFIKVATMGWMQNGFGSGNTGFSRR
			RARRQMGITISQV"
ORIGIN			
Query Match	91.6%;	Score 17.4;	DB 4; Length 1744;
Best Local Similarity	94.7%;	Pred. No. 2.6e+02;	
Matches	18; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	1	CTTTACTTCATAGCTTTTG	19
Db	267	CTTTACTTCATAGCTTTTG	285
RESULT 33			
SCYGR130C			
LOCUS			
DEFINITION			S.cerevisiae chromosome VII reading frame ORF YGR130c.
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
			Saccharomyces cerevisiae (baker's yeast)
			Saccharomyces cerevisiae
			Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
			Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE			
AUTHORS			van Dyck,L., Skala,J., de Wergifosse,P., Purnelle,B., Talla,E.,
			Nawrocki,A., Del Bino,S. and Goffeau,A.
JOURNAL			Unpublished
REFERENCE			
AUTHORS			2 (bases 1 to 1744)
			MIPS.
TITLE			Direct Submission
JOURNAL			Submitted (14-MAY-1996) Data collected by MIPS on behalf of the
			European yeast chromosome VII sequencing project. MIPS at the
			Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
			Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
FEATURES			
source			
			Location/Qualifiers
			1..1744
			/organism="Saccharomyces cerevisiae"
			/mol_type="genomic DNA"
			/db_xref="taxon:4932"
			/chromosome="VII"
			complement(744..754)
			/note="ARS-consensus"
			882..1406
			/note="unnamed protein product; ORF YGR131w"

Schackwitz,W.S., Sherwood,J.K., Sherwood,A.M., Leithauser,B.J. and
Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2004) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: NIEHS-SNPs, Environmental Genome
Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
(URL: <http://egp.gs.washington.edu>).

FEATURES	Location/Qualifiers	variation	variation
source	1..59195		3487
	/organism="Homo sapiens"		/frequency="0.01"
	/mol_type="genomic DNA"		/replace="a"
variation	119		3764
	/db_xref="taxon:9606"		/frequency="0.01"
variation	132		3772
	/replace="g"		/replace="c"
repeat_region	337..500		4121
	/rpt_family="MIR"		/frequency="0.01"
variation	453		4195
	/rpt_type=dispersed		/replace="c"
variation	499		4227
	/replace="c"		/frequency="0.01"
repeat_region	503..793		4243
	/rpt_family="Alu"		/replace="a"
variation	693..721		4555
	/frequency="0.46"		/gene="CHEK2"
variation	808		4594
	/replace="a"		/frequency="0.01"
repeat_region	1012..1260		4708
	/rpt_family="Alu"		/gene="CHEK2"
variation	1226		4746
	/rpt_type=dispersed		/replace="t"
repeat_region	1494..1791		4788
	/rpt_family="Alu"		/frequency="0.01"
variation	1520		4902
	/rpt_type=dispersed		/gene="CHEK2"
variation	1555		4905
	/replace="t"		/frequency="0.02"
variation	1767		4908
	/frequency="0.01"		/rpt_family="Alu"
variation	2230		5204
	/replace="g"		/rpt_type=dispersed
repeat_region	2254..2510		5276
	/rpt_family="L2"		/frequency="0.01"
repeat_region	2583..3015		5310
	/rpt_type=dispersed		/replace="a"
variation	3065		5413
	/rpt_family="ERV1"		/gene="CHEK2"
variation	3334		5413
	/replace="a"		/frequency="0.01"
variation	3434..3449		5413
	/frequency="0.99"		/replace="a"
	/replace="a"		/gene="CHEK2"


```

/replace="0.01"
5414
variation
/replace="t"
/replace="a"
5810
variation
/replace="a"
/replace="g"
5832
variation
/replace="a"
/replace="t"
5870.6157
repeat_region
/rpt_family="Alu"
/replace="t"
5878
variation
/replace="t"
/replace="g"
5915
variation
/replace="t"
/replace="t"
5984
variation
/replace="t"
/replace="t"
6158.6274
repeat_region
/rpt_family="ERVK"
/replace="t"
6177
variation
/replace="t"
/replace="g"
6224.8984
misc_feature
/replace="t"
/replace="g"
6275.6365
repeat_region

```

```

Query Match      91.6%; Score 17.4; DB 5; Length 59395;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 CTTTACTTCATAGCTCTTGG 19
Db 30742 CTTTACTTCATAGCTCTTGG 30760

```

```

RESULT 35
AC157840/c      108957 bp      DNA      linear      PRI 26-JAN-2006
LOCUS      AC157840
DEFINITION      Aotus nancymaae clone CH258-147H21, complete sequence.
ACCESSION      AC157840
VERSION      AC157840.3 GI:85719332
KEYWORDS      HTG.
SOURCE      Aotus nancymaae (Ma's night monkey)
ORGANISM

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
Cebidae; Aotinae; Aotus.
1 (bases 1 to 108957)
Antonellis,A., Bass,D., Benjamin,B., Bera,J., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
Franks,S., Fukeenko,T., Gestole,M., Greene,A., Guan,X., Gupta,J.,
Hunter,G., Hurler,B., Idol,J.R., Kwong,P., Latic,P., Larson,S.,
Lee-Tin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B.,
Montenayor,E.H., Masello,C., Maskeri,B., McDowell,J.,
Montemayor,C., Mullikin,J.C., Park,M., Prasad,A., Puri,O.,
Rantz,K., Reddix-Dugue,N., Sante,A., Schandler,K., Schueler,M.G.,
Sison,C., Stantripop,S., Taye,A., Thomas,J.W., Thomas,P.J.,

```

```

Tsipouri,V., Ung,L., Vogt,J.L., Wetherby,K.D., Withers,T.R.,
Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE      2 (bases 1 to 108957)
AUTHORS      Green,E.D.
TITLE      Direct Submission
JOURNAL      Submitted (02-MAR-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
REFERENCE      3 (bases 1 to 108957)
AUTHORS      Green,E.D.
TITLE      Direct Submission
JOURNAL      Submitted (07-APR-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
REFERENCE      4 (bases 1 to 108957)
AUTHORS      Green,E.D.
TITLE      Direct Submission
JOURNAL      Submitted (26-JAN-2006) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Jan 26, 2006 this sequence version replaced gi:62358351.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgri.nih.gov
----- Project Information
Center project name: hzv
Center clone name: 147H21

```

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

```

FEATURES
source

```

```

1.108957      Location/Qualifiers
/organism="Aotus nancymaae"
/mol_type="genomic DNA"
/db_xref="taxon:37293"
/clone="CH258-147H21"
/clone_lib="CH258"
/notes="BAC resource: http://bacpac.chori.org/"

```

```

misc_feature
1.97928

```

```

/notes="clone overlaps with GenBank Accession Number
AC151368.3 (nucleotides 85363-189563) clone CH258-251019
(center project name hzf)"

```

```

misc_feature
97935..>108957

```

```

/notes="this sequence is not the entire insert of clone
CH258-147H21; clone overlaps with GenBank Accession Number
AC151487.3 (nucleotides 1-56127) clone CH258-321B24
(center project name hzi); this annotated segment
represents overlap with nucleotides 1-11023 of AC151487.3"

```

```

ORIGIN

```

```

Query Match      91.6%; Score 17.4; DB 5; Length 108957;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 CTTTACTTCATAGCTCTTGG 19

```

```

Db 96005 CTTTACTTCATAGCTCTG 95987

```

```

RESULT 36
AP008215_067/c

```

WPCOMMENT

Sequence split into 227 fragments LOCUS AP008215 Accession AP008215

Fragment Name	Begin	End
AP008215_000	1	110000
AP008215_001	100001	210000
AP008215_002	200001	310000
AP008215_003	300001	410000
AP008215_004	400001	510000
AP008215_005	500001	610000
AP008215_006	600001	710000
AP008215_007	700001	810000
AP008215_008	800001	910000
AP008215_009	900001	1010000
AP008215_010	1000001	1110000
AP008215_011	1100001	1210000
AP008215_012	1200001	1310000
AP008215_013	1300001	1410000
AP008215_014	1400001	1510000
AP008215_015	1500001	1610000
AP008215_016	1600001	1710000
AP008215_017	1700001	1810000
AP008215_018	1800001	1910000
AP008215_019	1900001	2010000
AP008215_020	2000001	2110000
AP008215_021	2100001	2210000
AP008215_022	2200001	2310000
AP008215_023	2300001	2410000
AP008215_024	2400001	2510000
AP008215_025	2500001	2610000
AP008215_026	2600001	2710000
AP008215_027	2700001	2810000
AP008215_028	2800001	2910000
AP008215_029	2900001	3010000
AP008215_030	3000001	3110000
AP008215_031	3100001	3210000
AP008215_032	3200001	3310000
AP008215_033	3300001	3410000
AP008215_034	3400001	3510000
AP008215_035	3500001	3610000
AP008215_036	3600001	3710000
AP008215_037	3700001	3810000
AP008215_038	3800001	3910000
AP008215_039	3900001	4010000
AP008215_040	4000001	4110000
AP008215_041	4100001	4210000
AP008215_042	4200001	4310000
AP008215_043	4300001	4410000
AP008215_044	4400001	4510000
AP008215_045	4500001	4610000
AP008215_046	4600001	4710000
AP008215_047	4700001	4810000
AP008215_048	4800001	4910000
AP008215_049	4900001	5010000
AP008215_050	5000001	5110000
AP008215_051	5100001	5210000
AP008215_052	5200001	5310000
AP008215_053	5300001	5410000
AP008215_054	5400001	5510000
AP008215_055	5500001	5610000
AP008215_056	5600001	5710000
AP008215_057	5700001	5810000
AP008215_058	5800001	5910000
AP008215_059	5900001	6010000
AP008215_060	6000001	6110000
AP008215_061	6100001	6210000
AP008215_062	6200001	6310000
AP008215_063	6300001	6410000
AP008215_064	6400001	6510000
AP008215_065	6500001	6610000
AP008215_066	6600001	6710000
AP008215_067	6700001	6810000
AP008215_068	6800001	6910000
AP008215_069	6900001	7010000
AP008215_070	7000001	7110000
AP008215_071	7100001	7210000
AP008215_072	7200001	7310000
AP008215_073	7300001	7410000
AP008215_074	7400001	7510000
AP008215_075	7500001	7610000
AP008215_076	7600001	7710000
AP008215_077	7700001	7810000
AP008215_078	7800001	7910000
AP008215_079	7900001	8010000
AP008215_080	8000001	8110000
AP008215_081	8100001	8210000
AP008215_082	8200001	8310000
AP008215_083	8300001	8410000
AP008215_084	8400001	8510000
AP008215_085	8500001	8610000
AP008215_086	8600001	8710000
AP008215_087	8700001	8810000
AP008215_088	8800001	8910000
AP008215_089	8900001	9010000
AP008215_090	9000001	9110000
AP008215_091	9100001	9210000
AP008215_092	9200001	9310000
AP008215_093	9300001	9410000
AP008215_094	9400001	9510000
AP008215_095	9500001	9610000
AP008215_096	9600001	9710000
AP008215_097	9700001	9810000
AP008215_098	9800001	9910000
AP008215_099	9900001	10010000
AP008215_100	10000001	10110000
AP008215_101	10100001	10210000
AP008215_102	10200001	10310000
AP008215_103	10300001	10410000
AP008215_104	10400001	10510000
AP008215_105	10500001	10610000
AP008215_106	10600001	10710000
AP008215_107	10700001	10810000
AP008215_108	10800001	10910000
AP008215_109	10900001	11010000
AP008215_110	11000001	11110000
AP008215_111	11100001	11210000
AP008215_112	11200001	11310000
AP008215_113	11300001	11410000
AP008215_114	11400001	11510000
AP008215_115	11500001	11610000
AP008215_116	11600001	11710000
AP008215_117	11700001	11810000
AP008215_118	11800001	11910000
AP008215_119	11900001	12010000
AP008215_120	12000001	12110000
AP008215_121	12100001	12210000
AP008215_122	12200001	12310000
AP008215_123	12300001	12410000
AP008215_124	12400001	12510000
AP008215_125	12500001	12610000
AP008215_126	12600001	12710000
AP008215_127	12700001	12810000
AP008215_128	12800001	12910000
AP008215_129	12900001	13010000
AP008215_130	13000001	13110000
AP008215_131	13100001	13210000
AP008215_132	13200001	13310000
AP008215_133	13300001	13410000
AP008215_134	13400001	13510000
AP008215_135	13500001	13610000
AP008215_136	13600001	13710000
AP008215_137	13700001	13810000
AP008215_138	13800001	13910000
AP008215_139	13900001	14010000
AP008215_140	14000001	14110000
AP008215_141	14100001	14210000
AP008215_142	14200001	14310000

AP008215_143	14300001	14410000	AP008207_004	400001	510000
AP008215_144	14400001	14510000	AP008207_005	500001	610000
AP008215_145	14500001	14610000	AP008207_006	600001	710000
AP008215_146	14600001	14710000	AP008207_007	700001	810000
AP008215_147	14700001	14810000	AP008207_008	800001	910000
AP008215_148	14800001	14910000	AP008207_009	900001	1010000
AP008215_149	14900001	15010000	AP008207_010	1000001	1110000
AP008215_150	15000001	15110000	AP008207_011	1100001	1210000
AP008215_151	15100001	15210000	AP008207_012	1200001	1310000
AP008215_152	15200001	15310000	AP008207_013	1300001	1410000
AP008215_153	15300001	15410000	AP008207_014	1400001	1510000
AP008215_154	15400001	15510000	AP008207_015	1500001	1610000
AP008215_155	15500001	15610000	AP008207_016	1600001	1710000
AP008215_156	15600001	15710000	AP008207_017	1700001	1810000
AP008215_157	15700001	15810000	AP008207_018	1800001	1910000
AP008215_158	15800001	15910000	AP008207_019	1900001	2010000
AP008215_159	15900001	16010000	AP008207_020	2000001	2110000
AP008215_160	16000001	16110000	AP008207_021	2100001	2210000
AP008215_161	16100001	16210000	AP008207_022	2200001	2310000
AP008215_162	16200001	16310000	AP008207_023	2300001	2410000
AP008215_163	16300001	16410000	AP008207_024	2400001	2510000
AP008215_164	16400001	16510000	AP008207_025	2500001	2610000
AP008215_165	16500001	16610000	AP008207_026	2600001	2710000
AP008215_166	16600001	16710000	AP008207_027	2700001	2810000
AP008215_167	16700001	16810000	AP008207_028	2800001	2910000
AP008215_168	16800001	16910000	AP008207_029	2900001	3010000
AP008215_169	16900001	17010000	AP008207_030	3000001	3110000
AP008215_170	17000001	17110000	AP008207_031	3100001	3210000
AP008215_171	17100001	17210000	AP008207_032	3200001	3310000
AP008215_172	17200001	17310000	AP008207_033	3300001	3410000
AP008215_173	17300001	17410000	AP008207_034	3400001	3510000
AP008215_174	17400001	17510000	AP008207_035	3500001	3610000
AP008215_175	17500001	17610000	AP008207_036	3600001	3710000
AP008215_176	17600001	17710000	AP008207_037	3700001	3810000
AP008215_177	17700001	17810000	AP008207_038	3800001	3910000
AP008215_178	17800001	17910000	AP008207_039	3900001	4010000
AP008215_179	17900001	18010000	AP008207_040	4000001	4110000
AP008215_180	18000001	18110000	AP008207_041	4100001	4210000
AP008215_181	18100001	18210000	AP008207_042	4200001	4310000
AP008215_182	18200001	18310000	AP008207_043	4300001	4410000
AP008215_183	18300001	18410000	AP008207_044	4400001	4510000
AP008215_184	18400001	18510000	AP008207_045	4500001	4610000
AP008215_185	18500001	18610000	AP008207_046	4600001	4710000
AP008215_186	18600001	18710000	AP008207_047	4700001	4810000
AP008215_187	18700001	18810000	AP008207_048	4800001	4910000
AP008215_188	18800001	18910000	AP008207_049	4900001	5010000
AP008215_189	18900001	19010000	AP008207_050	5000001	5110000
AP008215_190	19000001	19110000	AP008207_051	5100001	5210000
AP008215_191	19100001	19210000	AP008207_052	5200001	5310000
AP008215_192	19200001	19310000	AP008207_053	5300001	5410000
AP008215_193	19300001	19410000	AP008207_054	5400001	5510000
AP008215_194	19400001	19510000	AP008207_055	5500001	5610000
AP008215_195	19500001	19610000	AP008207_056	5600001	5710000
AP008215_196	19600001	19710000	AP008207_057	5700001	5810000
			AP008207_058	5800001	5910000
			AP008207_059	5900001	6010000
			AP008207_060	6000001	6110000
			AP008207_061	6100001	6210000
			AP008207_062	6200001	6310000
			AP008207_063	6300001	6410000
			AP008207_064	6400001	6510000
			AP008207_065	6500001	6610000
			AP008207_066	6600001	6710000
			AP008207_067	6700001	6810000
			AP008207_068	6800001	6910000
			AP008207_069	6900001	7010000
			AP008207_070	7000001	7110000
			AP008207_071	7100001	7210000
			AP008207_072	7200001	7310000
			AP008207_073	7300001	7410000
			AP008207_074	7400001	7510000
			AP008207_075	7500001	7610000
			AP008207_076	7600001	7710000

Query Match Best Local Similarity 91.6%; Score 17.4; DB 4; Length 110000;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
 Db 67251 CTTTACTTCATAGCTTCG 67233

RESULT 37
 WPCOMMENT
 Sequence split into 433 fragments LOCUS AP008207 Accession AP008207
 Fragment Name Begin End
 AP008207_000 1 110000
 AP008207_001 100001 210000
 AP008207_002 200001 310000
 AP008207_003 300001 410000

AP008207_077 7700001 7810000
AP008207_078 7800001 7910000
AP008207_079 7900001 8010000
AP008207_080 8000001 8110000
AP008207_081 8100001 8210000
AP008207_082 8200001 8310000
AP008207_083 8300001 8410000
AP008207_084 8400001 8510000
AP008207_085 8500001 8610000
AP008207_086 8600001 8710000
AP008207_087 8700001 8810000
AP008207_088 8800001 8910000
AP008207_089 8900001 9010000
AP008207_090 9000001 9110000
AP008207_091 9100001 9210000
AP008207_092 9200001 9310000
AP008207_093 9300001 9410000
AP008207_094 9400001 9510000
AP008207_095 9500001 9610000
AP008207_096 9600001 9710000
AP008207_097 9700001 9810000
AP008207_098 9800001 9910000
AP008207_099 9900001 10010000
AP008207_100 10000001 10110000
AP008207_101 10100001 10210000
AP008207_102 10200001 10310000
AP008207_103 10300001 10410000
AP008207_104 10400001 10510000
AP008207_105 10500001 10610000
AP008207_106 10600001 10710000
AP008207_107 10700001 10810000
AP008207_108 10800001 10910000
AP008207_109 10900001 11010000
AP008207_110 11000001 11110000
AP008207_111 111000001 11210000
AP008207_112 11200001 11310000
AP008207_113 11300001 11410000
AP008207_114 11400001 11510000
AP008207_115 11500001 11610000
AP008207_116 11600001 11710000
AP008207_117 11700001 11810000
AP008207_118 11800001 11910000
AP008207_119 11900001 12010000
AP008207_120 12000001 12110000
AP008207_121 12100001 12210000
AP008207_122 12200001 12310000
AP008207_123 12300001 12410000
AP008207_124 12400001 12510000
AP008207_125 12500001 12610000
AP008207_126 12600001 12710000
AP008207_127 12700001 12810000
AP008207_128 12800001 12910000
AP008207_129 12900001 13010000
AP008207_130 13000001 13110000
AP008207_131 13100001 13210000
AP008207_132 13200001 13310000
AP008207_133 13300001 13410000
AP008207_134 13400001 13510000
AP008207_135 13500001 13610000
AP008207_136 13600001 13710000
AP008207_137 13700001 13810000
AP008207_138 13800001 13910000
AP008207_139 13900001 14010000
AP008207_140 14000001 14110000
AP008207_141 14100001 14210000
AP008207_142 14200001 14310000
AP008207_143 14300001 14410000
AP008207_144 14400001 14510000
AP008207_145 14500001 14610000
AP008207_146 14600001 14710000
AP008207_147 14700001 14810000
AP008207_148 14800001 14910000
AP008207_149 14900001 15010000

AP008207_150 15000001 15110000
AP008207_151 15100001 15210000
AP008207_152 15200001 15310000
AP008207_153 15300001 15410000
AP008207_154 15400001 15510000
AP008207_155 15500001 15610000
AP008207_156 15600001 15710000
AP008207_157 15700001 15810000
AP008207_158 15800001 15910000
AP008207_159 15900001 16010000
AP008207_160 16000001 16110000
AP008207_161 16100001 16210000
AP008207_162 16200001 16310000
AP008207_163 16300001 16410000
AP008207_164 16400001 16510000
AP008207_165 16500001 16610000
AP008207_166 16600001 16710000
AP008207_167 16700001 16810000
AP008207_168 16800001 16910000
AP008207_169 16900001 17010000
AP008207_170 17000001 17110000
AP008207_171 17100001 17210000
AP008207_172 17200001 17310000
AP008207_173 17300001 17410000
AP008207_174 17400001 17510000
AP008207_175 17500001 17610000
AP008207_176 17600001 17710000
AP008207_177 17700001 17810000
AP008207_178 17800001 17910000
AP008207_179 17900001 18010000
AP008207_180 18000001 18110000
AP008207_181 18100001 18210000
AP008207_182 18200001 18310000
AP008207_183 18300001 18410000
AP008207_184 18400001 18510000
AP008207_185 18500001 18610000
AP008207_186 18600001 18710000
AP008207_187 18700001 18810000
AP008207_188 18800001 18910000
AP008207_189 18900001 19010000
AP008207_190 19000001 19110000
AP008207_191 19100001 19210000
AP008207_192 19200001 19310000
AP008207_193 19300001 19410000
AP008207_194 19400001 19510000
AP008207_195 19500001 19610000
AP008207_196 19600001 19710000

Query Match 91.6%; Score 17.4; DB 4; Length 110000;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 CTTTACTTCATAGTCTTTG 19
||| ||||| ||||| |||||
Db 20335 CTTTCTTCATAGTCTTTG 20353

RESULT 38
CP000155_24/c
WPCOMMENT

Sequence split into 73 fragments LOCUS CP000155 Accession CP000155
Fragment Name Begin End
CP000155_00 1 110000
CP000155_01 100001 210000
CP000155_02 200001 310000
CP000155_03 300001 410000
CP000155_04 400001 510000
CP000155_05 500001 610000
CP000155_06 600001 710000
CP000155_07 700001 810000
CP000155_08 800001 910000
CP000155_09 900001 1010000
CP000155_10 1000001 1110000

```
CP000155-11 1100001 1210000
CP000155-12 1200001 1310000
CP000155-13 1300001 1410000
CP000155-14 1400001 1510000
CP000155-15 1500001 1610000
CP000155-16 1600001 1710000
CP000155-17 1700001 1810000
CP000155-18 1800001 1910000
CP000155-19 1900001 2010000
CP000155-20 2000001 2110000
CP000155-21 2100001 2210000
CP000155-22 2200001 2310000
CP000155-23 2300001 2410000
CP000155-24 2400001 2510000
CP000155-25 2500001 2610000
CP000155-26 2600001 2710000
CP000155-27 2700001 2810000
CP000155-28 2800001 2910000
CP000155-29 2900001 3010000
CP000155-30 3000001 3110000
CP000155-31 3100001 3210000
CP000155-32 3200001 3310000
CP000155-33 3300001 3410000
CP000155-34 3400001 3510000
CP000155-35 3500001 3610000
CP000155-36 3600001 3710000
CP000155-37 3700001 3810000
CP000155-38 3800001 3910000
CP000155-39 3900001 4010000
CP000155-40 4000001 4110000
CP000155-41 4100001 4210000
CP000155-42 4200001 4310000
CP000155-43 4300001 4410000
CP000155-44 4400001 4510000
CP000155-45 4500001 4610000
CP000155-46 4600001 4710000
CP000155-47 4700001 4810000
CP000155-48 4800001 4910000
CP000155-49 4900001 5010000
CP000155-50 5000001 5110000
CP000155-51 5100001 5210000
CP000155-52 5200001 5310000
CP000155-53 5300001 5410000
CP000155-54 5400001 5510000
CP000155-55 5500001 5610000
CP000155-56 5600001 5710000
CP000155-57 5700001 5810000
CP000155-58 5800001 5910000
CP000155-59 5900001 6010000
CP000155-60 6000001 6110000
CP000155-61 6100001 6210000
CP000155-62 6200001 6310000
CP000155-63 6300001 6410000
CP000155-64 6400001 6510000
CP000155-65 6500001 6610000
CP000155-66 6600001 6710000
CP000155-67 6700001 6810000
CP000155-68 6800001 6910000
CP000155-69 6900001 7010000
CP000155-70 7000001 7110000
CP000155-71 7100001 7210000
CP000155-72 7200001 7215267

Continuation (25 of 73) of CP000155 from base 2400001 (CP000155 Hahella chejuensis KCTC

Query Match 91.6%; Score 17.4; DB 15; Length 110000;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
| | | | | | | | | | | | | | | | | | | | |
Db 60999 CTTTACTTCATAGCTTTG 60881

RESULT 39
AP001818/c 116321 bp DNA linear PRI 24-OCT-2000
LOCUS Homo sapiens genomic DNA, chromosome 4q22-q24, clone:2222M9,
DEFINITION complete sequence.
ACCESSION AP001818 GI:10998160
VERSION AP001818.2
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (sites)
Tsai,S.F.
TITLE Genome Sequencing of the Chromosome 4q Region Implicated in Human
Hepatocellular Carcinoma Pathogenesis 1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 116321)
Tsai,S.F.
AUTHORS Direct Submission
TITLE Submitted (14-APR-2000) Shih-Feng Tsai, National Yang-Ming
JOURNAL University, Institute of Genetics; 155 Li-Rong St. Section 2,
Peitou, Taipei, Taiwan 11221, Republic of China
(E-mail:ymptsai@ym.edu.tw, URL:http://genome.ym.edu.tw/,
Tel:886-2-28267043, Fax:886-2-28264930)
On Oct 24, 2000 this sequence version replaced gi:7594883.
COMMENT Project AP
Quality: the expected Phred/Phrap calculated error rate (per 10 kb)
is 0.0007; Estimated total number of errors is 0.0087.
FEATURES
source
1. .116321
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"
/chromosome="4"
/map="4q22-q24"
/clone="2222M9"

ORIGIN
Query Match 91.6%; Score 17.4; DB 5; Length 116321;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
| | | | | | | | | | | | | | | | | | | | |
Db 79498 CATTACTTCATAGCTTTG 79480

RESULT 40
AC150029/c 121462 bp DNA linear HTG 05-AUG-2004
LOCUS Sorex araneus clone SA_Ba-362A8, WORKING DRAFT SEQUENCE, 2 ordered
DEFINITION pieces.
ACCESSION AC150029
VERSION AC150029.2 GI:50979356
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Sorex araneus (European shrew)
ORGANISM Sorex araneus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Insectivora; Soricidae;
Soricinae; Sorex.
REFERENCE 1 (bases 1 to 121462)
Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
Karlsn,B., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-lin,S.-Q.,
Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
Masello,C., Maskeri,B., McDowell,J., Mullikin,J.C., Pauguiran,C.,
Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddix-Dugue,N.,
Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantripop,S.,
Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D.,
```

```

Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 121462)
Green, E.D.
Direct Submission
Submitted (30-JUN-2004) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 121462)
Green, E.D.
Direct Submission
Submitted (05-AUG-2004) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Aug 5, 2004 this sequence version replaced gi:49457917.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: ggc
Center clone name: 362A08

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 121178 bases at least Q40
Consensus quality: 121304 bases at least Q30
Consensus quality: 121341 bases at least Q20
Insert size: 125000; agarose-fp
Quality coverage: 121362; sum-of-contigs
Quality coverage: 8.26x in Q20 bases; agarose-fp
Quality coverage: 8.51x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 25692: contig of 25692 bp in length
* 25693 25792: gap of unknown length
* 25793 121462: contig of 95670 bp in length.
Location/Qualifiers
1. 121462
/organism="Sorex araneus"
/mol_type="genomic DNA"
/db_xref="taxon:42254"
/clone="SA Ba-362A8"
/clone_lib="SA_Ba"
misc_feature 1. 25692
/note="assembly_fragment
clone_end:r7
vector_side:left"
25693. 25792
/estimated_length=unknown
misc_feature 25793. 121462

```

```

/note="assembly_fragment
clone_end:SP6
vector_side:right"

Query Match 91.6%; Score 17.4; DB 12; Length 121462;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 51546 CTTTACTTCATAGTATTTG 51528
|||||

RESULT 41
AC167325 121581 bp DNA linear HTG 24-NOV-2005
LOCUS Loxodonta africana clone VMRC15-443K8, WORKING DRAFT SEQUENCE, 8
DEFINITION ordered pieces.
ACCESSION AC167325
VERSION AC167325.2 GI:82654480
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Loxodonta africana (African savanna elephant)
ORGANISM Loxodonta africana
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Afrotheria; Proboscidea; Elephantidae;
Loxodonta.
1 (bases 1 to 121581)
Antonellis, A., Ayele, K., Bass, D., Benjamin, B., Bera, J.,
Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G.,
Coleman, H., Franks, S., Fukukenko, T., Gestole, M., Greene, A., Guan, X.,
Gupta, J., Gurson, N., Haghighi, P., Han, E., Han, J., Hansen, N.,
Ho, S.-L., Hu, P., Hunter, G., Hurie, B., Idol, J.R., Kwong, P.,
Laric, P., Larson, S., Lee-Lin, S.-Q., Legaspi, R., Madden, M.,
Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C., Park, M., Portnoy, M.E.,
McDowell, J., Montemayor, C., Mullikin, J.C., Park, M., Portnoy, M.E.,
Prasad, A., Puri, O., Rantz, K., Reddix-Dugue, N., Sante, A.,
Schandler, K., Schueler, M.G., Sison, C., Stantripop, S., Tave, A.,
Wetherby, K.D., Withers, T.R., Young, A., Ung, L., Vogt, J.L.,
Thomas, J.W., Thomas, P.J., Tsipouri, V., Ung, L., Vogt, J.L.,
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 121581)
Green, E.D.
Direct Submission
Submitted (24-AUG-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
3 (bases 1 to 121581)
Green, E.D.
Direct Submission
Submitted (24-NOV-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Nov 24, 2005 this sequence version replaced gi:73745553.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: jrf
Center clone name: 443K08

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is generally based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated

```

```

with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 119217 bases at least Q40
Consensus quality: 120157 bases at least Q30
Consensus quality: 120697 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 120881; sum-of-contigs
Quality coverage: 9.21x in Q20 bases; agarose-fp
Quality coverage: 11.27x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 47612: contig of 47612 bp in length
* 47613: gap of unknown length
* 47713: contig of 8797 bp in length
* 56510: gap of unknown length
* 56609: gap of unknown length
* 56610: contig of 3067 bp in length
* 59676: gap of unknown length
* 59677: contig of 13480 bp in length
* 73256: gap of unknown length
* 73257: contig of 9293 bp in length
* 73257: gap of unknown length
* 82649: contig of 2795 bp in length
* 82749: gap of unknown length
* 82750: contig of 21225 bp in length
* 85644: gap of unknown length
* 85645: contig of 14612 bp in length.
* 85645: gap of unknown length
* 106870: contig of 14612 bp in length.
* 106970: Location/Qualifiers
          1..121581
            /organism="Loxodonta africana"
            /mol_type="genomic DNA"
            /db_xref="taxon:9785"
            /clone="VMRC15-443K8"
            /clone_lib="VMRC15"
            /note="BAC resource: http://bacpac.chori.org/"
            1..47612
              /note="assembly fragment
              missing SP6 clone_end on 5' of insert"
              44142..121581
                /note="clone overlaps with GenBank Accession Number
                AC163963 clone VMRC15-243L21 (center project name jrg)"
                47613..47712
                  /estimated_length=unknown
                  47713..56509
                    /note="assembly_fragment"
                    56510..56609
                      /estimated_length=unknown
                      56610..59676
                        /note="assembly_fragment"
                        59677..59776
                          /estimated_length=unknown
                          59777..73256
                            /note="assembly_fragment"
                            73257..73356
                              /estimated_length=unknown
                              73357..82649
                                /note="assembly_fragment"
                                82650..82749
                                  /estimated_length=unknown
                                  82750..85544
                                    /note="assembly_fragment"
                                    85545..85644
                                      /estimated_length=unknown
                                      /estimated_length=unknown

```

```

misc_feature      85645..106869
                  /note="assembly_fragment"
gap              106870..106969
                  /estimated_length=unknown
misc_feature      106970..121581
                  /note="assembly_fragment
                  clone_end:17
                  vector_side:right"
ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 121581;
Best Local Similarity 94.7%; Pred.No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
      |||||
Db 23777 CTTTACTTCATAGTCATTG 23795
      |||||

RESULT 42
HSBA436C9/c
LOCUS            HSBA436C9      124337 bp      DNA      linear      PRI 18-MAY-2005
DEFINITION      Human DNA sequence from clone RP11-436C9 on chromosome 22, complete
                  sequence.
ACCESSION      AL121825
VERSION      AL121825.19      GI:6911603
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Homnidae; Homo.
REFERENCE      1 (bases 1 to 124337)
AUTHORS      Clark,G.
TITLE      Direct Submission
JOURNAL      Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
                Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT      Clone requests: clonerequest@sanger.ac.uk
                On Feb 7, 2000 this sequence version replaced gi:6469587.
                The following abbreviations are used to associate primary accession
                numbers given in the feature table with their source databases:
                En.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
                on the WORMPEP database can be found at
                http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
                was generated from part of bacterial clone contigs of human
                chromosome 22, constructed by the Sanger Centre Chromosome 22
                Mapping Group. Further information can be found at
                http://www.sanger.ac.uk/HGP/Chr22
                RP11-436C9 is from the library RPCI-11.2 constructed by the group
                of Pieter de Jong. For further details see
                http://www.chori.org/bacpac/home.htm
                VECTOT: pBACe3.6
                ----- Genome Center
                Center: Wellcome Trust Sanger Institute
                Center code: SC
                Web site: http://www.sanger.ac.uk
                Contact: vegas@sanger.ac.uk
                -----
                This sequence was finished as follows unless otherwise noted: all
                regions were either double-stranded or sequenced with an alternate
                chemistry or covered by high quality data (i.e., phred quality >=
                30); an attempt was made to resolve all sequencing problems, such
                as compressions and repeats; all regions were covered by at least
                one subclone; and the assembly was confirmed by restriction digest,
                except on the rare occasion of the clone being a YAC.
                Location/Qualifiers
                1..124337
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="22"
                  /clone="RP11-436C9"
                  /clone_lib="RPCI-11.2"
FEATURES
source

```



```
100 misc_feature
/Note="Clone_right_end: RP11-329J7"
6496 6642
/Note="Single clone region. pUC subcloned to give 5X
coverage over region."
join(complement(37948..88208),
complement(AL023281.1:85702..96147),
complement(AL035397.4:30093..30600),
complement(AL050313.6:101..1143),
complement(AL035453.4:18624..31904),
complement(AL035358.1:17058..69410))
/locus_tag="RP3-477H23.1-001"
join(complement(88018..88208),complement(37948..38226),
complement(AL023281.1:96000..96147),
complement(AL023281.1:87084..87356),
complement(AL023281.1:85702..85832),
complement(AL035397.4:30093..30600),
complement(AL050313.6:101..1143),
complement(AL035453.4:18624..31904),
complement(AL035453.4:29823..30346),
complement(AL035453.4:25715..25824),
complement(AL035453.4:23459..23588),
complement(AL035453.4:20734..20952),
complement(AL035453.4:18624..18789),
complement(AL03538.1:69270..69410),
complement(AL03538.1:53292..53436),
complement(AL03538.1:40418..40597),
complement(AL03538.1:37584..38304),
complement(AL03538.1:35179..35303),
complement(AL03538.1:32331..32562),
complement(AL03538.1:31631..31707),
complement(AL03538.1:29833..29856),
complement(AL03538.1:28922..29051),
complement(AL03538.1:24293..24400),
complement(AL03538.1:17058..22895))
/locus_tag="RP3-477H23.1-001"
join(complement(88018..88119),complement(37948..38226),
complement(AL023281.1:96000..96147),
complement(AL023281.1:87084..87356),
complement(AL023281.1:85702..85832),
complement(AL035397.4:30093..30600),
complement(AL050313.6:101..1143),
complement(AL035453.4:18624..31904),
complement(AL035453.4:29823..30346),
complement(AL035453.4:25715..25824),
complement(AL035453.4:23459..23588),
complement(AL035453.4:20734..20952),
complement(AL035453.4:18624..18789),
complement(AL03538.1:69270..69410),
complement(AL03538.1:53292..53436),
complement(AL03538.1:40418..40597),
complement(AL03538.1:37584..38304),
complement(AL03538.1:35179..35303),
complement(AL03538.1:32331..32562),
complement(AL03538.1:31631..31707),
complement(AL03538.1:29833..29856),
complement(AL03538.1:28922..29051),
complement(AL03538.1:24293..24400),
complement(AL03538.1:17058..22895))
/locus_tag="RP3-477H23.1-001"
/standard_name="OTTHUMP0000028696"
/codon_start=1
/protein_id="CAH73824.1"
/db_xref="GI:5560890"
/db_xref="InterPro:IPR001440"
/db_xref="InterPro:IPR011595"
/db_xref="InterPro:IPR011716"
/db_xref="InterPro:IPR011717"
/db_xref="UniProtKB/TrEMBL:Q5W189"
/translaton="MEQSPPPAPEPTGPTPARRRRREPPSPASAPIPLFGADTTG
QRSPDGVLSKAEFVKVQSNQACHDGFHTAIVLYNEALAVDPQCILYSNRSAAY
MKIQYDKALDDAIKARLLNPKPKYFRQGVALLQGRHADALAAFAAGLAQDPKSL
QLLVGVWEAAKMSPMRDSLEPTYQLQKMKLDKSPFVVSVVVGQELLTAGHHGASVVV
```

```
LEALKIGTCSLKLGRSVFSGALSSAYMSLGNTEKSTGYMOODLDVAKTLGDTGECRA
HGNLGSFAFKNGYREALTNHRHOLVLA MKLKDREAAASSLSSLGHVHTAIGDYNAL
ASHQCVLLAKQSKDELSEARELGNMGAVYIANGDFENAVOCHQEHQHLKIAKDLGNKRE
EARYSNLGSAYHYRRNFDKXMSYHNYVLELAQELMEKAEMRAYAGLGHARCNQDL
EYAKQHEHQGLGAEDUDKRAEGRASSNLGIITHQMGDYDTDLAKHKTUHCIAQELS
DYAAQGRAYGNMAYNALGYMDOAVYRHOELIQTSMEDNDRASQASTHGLNAVAYQA
LGADRALQHOYHNLINARELURDIQSEARALSNLGNFCHSRGEYQAPAYEYQLRLA
PDLQMEGEGKVCNLYGAYHCLNGYQAVKYQEYQDLAKDLHDKLQAKAYCNLGL
AFKALLNKAECOCKYLLSQAOLNNSQAKFRALGNLGDIFICKDINGAIKPYEQQ
LGLAHQVDRLEASAYALGLTAYMTOKYDKALGYHTOLEVLEQSLDLPGECPRAHG
HLAAYMALGKTYTNAFKYEBQDLGQKLDKPSLEAQYNGMGITKMNWVWEAIGY
FEQOALMLQQLSGNESVLDGRAYNGLDYCEALGDYEEAIKYQEYLSVAQSLNRMQ
DOKAYRGLNGHRAMGSLQALVCFEKLIVVAHELGEAFNAKQAYGELSLHSLQNLN
YBOAISCLERQLNITARDKDRALSDAACGLGVYQOMGYEDYALQHOLDIQAETEH
NNPTCOGRAYNGLTYESLGTFRERAVYQEOHLSIAAOMNDLAAKTYSYSLGRTHH
ALQNSQAVMYLQEGRLAEQLGREREDEAKIRHGLGLSLWASGNLEEAHQOLYRASAL
FETIRHEAUSTDYKSLFDLQTSYQALQRLVLSLGHHDALAEGRTRAFADLL
VERQTQGDSDPYSPVTIDQILEMVGORGLVLYSLAAGLYLWLLAPGAGIYKVFHE
EAGIDMQOFEEMNNKLSVTDPTGFLRMVRRNNLFNRSCOSMTSLFSNTVSPQDGT
SSLPROSSFAKPLRALYDLLIAPMEGLMHSQSPVGRHRLQILVLEGLYLIPLFAL
LKSSSNEYLYERFGLLAVPSIRSLSVQSKSHLRKNPTTYSSTMSMAAVIGNPKLPISA
VMRWLMGPMPSAEEAYMVSELLGQCLPLVSGVATKRVMSALTQAEVCVHFATHLSWK
LSALVLTPSKVDGNPSSKSGFHEPTTIPESLRVQDDASDGESI SDCPLQLLLTAAD
VLQDLPLVKLVLGSSQESNSKVTADGVIALTRAFLAAGACQVLVSLWMPVPAASKMF
IHAFYSSLLNGLKASALGRAMKVVQSKAFSPHSPNAGFMILGSDVKLNSPSSLIGQ
ALTEILOHPRARDALRVLLHVEKSLQRLONGORNAMYTSSQSVENKGGIPGQWAL
LTVAGFRDLPPTSGLPAAVFPPTSDGRLQOCSSTLQSLGLLGNPALQALCKLITAS
ETUEQLISRAVKNNVLMHQLVOLQAGEKEQDLASAPIQVSIQVLRWLRLPGCHEFLA
ALQFDLCEVQEEVILTKQANRRTVHFALQSLLSLFDSTELPKRLSLSSSLESLE
ASAQSDVSNALPLGYQRPFPSTGADSIADSAISVYLSIASSMSFVSKPRGSEGG
PGQRQDHRKSNAYLQBSTLPSRLPOTRPNAGNDEEYEGFSIISNEPLATYOENR
NTCFSPDHKOPQCTAGGMVSVSSKGSISTPNSPVKMTLIPSPNSPKQVKGLASSD
TGSDQSTETDSTVKSQESNPKLDQELAQKILESTQSHLJAVERLQSGGQVSKS
NNEDGQVQSPSTAVFRASSTAFSPVLSHQKSPSPVTKVPPKARSSSLPKVSSG
YSPPTTSEMISKDPSQSGRPSGCDSTQSLDPLFKLYPSPSPSAHYSKSPRNM
SPSGHSGPAGSPASPALSYSAGSARSPADAPDIDKLKMAAIDEKVQAYHNLKMF
QSTPQHSHTGPMKIFRGAFTMTSKRDVLSLNLSPRNKKEEGVDKLELSLQOHD
GAPPKAPNGHWRTETTSLSGLPLPAGPAPATAPARPLRPSGNGKFLSPGRFFPSK
C"
43142..43210
/Note="Single clone region. pUC subcloned to give 3x
coverage over region."
52224..52271
/Note="Weak data"
69442..69505
/Note="Weak data"
join(complement(AL117330.6:3554..26647),
complement(96139..120413))
/locus_tag="RP11-436C9.1-001"
join(complement(AL117330.6:25928..26647),
complement(AL117330.6:18562..18886),
complement(AL117330.6:9402..9526),
complement(AL117330.6:9136..9283),
complement(AL117330.6:3554..3644),
complement(120305..120413),complement(118402..118455),
complement(111901..111962),complement(108234..108333),
complement(105297..105383),complement(104106..104269),
complement(103523..103638),complement(102428..102513),
complement(97531..97611),complement(96139..96382))
/locus_tag="RP11-436C9.1-001"
join(complement(AL117330.6:18562..18880),
complement(AL117330.6:9402..9526),
complement(AL117330.6:9136..9283),
complement(AL117330.6:3554..3644),
complement(120305..120413),complement(118402..118455),
complement(111901..111962),complement(108234..108333),
complement(105297..105383),complement(104106..104269),
complement(103523..103638),complement(102428..102513),
complement(97531..97611),complement(96293..96382))
/locus_tag="RP11-436C9.1-001"
/standard_name="OTTHUMP0000028871"
/codon_start=1
```

```

/protein_id="CAH73823.1"
/db_xref="GI:5560889"

Query Match      91.6%; Score 17.4; DB 5; Length 124337;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTTG 19
|||||
Db 123733 CTTTACTTCATAGCTTTTG 123715

RESULT 43
AC011354/c
LOCUS      133042 bp DNA linear PRI 27-JUN-2001
DEFINITION Homo sapiens chromosome 5 clone CTC-354F19, complete sequence.
ACCESSION  AC011354
VERSION     AC011354.4 GI:14572125
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chordata; Euarthontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE  1 (bases 1 to 133042)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            TITLE Direct Submission
REFERENCE  2 (bases 1 to 133042)
            DOE Joint Genome Institute.
            TITLE Direct Submission
REFERENCE  3 (bases 1 to 133042)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            TITLE Direct Submission
            JOURNAL Submitted (27-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE  3 (bases 1 to 133042)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            TITLE Direct Submission
            JOURNAL Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE  3 (bases 1 to 133042)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            TITLE Direct Submission
            JOURNAL Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Nov 27, 2001 this sequence version replaced gi:13699555.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
            Location/Qualifiers
            1..133042
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="5"
               /clones="CTC-354F19"

FEATURES             source
     source           1..133042
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="5"
                        /clones="CTC-354F19"

ORIGIN
Query Match      91.6%; Score 17.4; DB 5; Length 133042;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTTG 19
|||||
Db 95584 CTTTACTTCATAGCTTTTG 95566

RESULT 44
AC011334/c
LOCUS      134413 bp DNA linear PRI 27-NOV-2001
DEFINITION Homo sapiens chromosome 5 clone CTC-235N17, complete sequence.
ACCESSION  AC011334
VERSION     AC011334.5 GI:17105283
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chordata; Euarthontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE  1 (bases 1 to 134413)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            TITLE Direct Submission
REFERENCE  2 (bases 1 to 134413)
            DOE Joint Genome Institute.
            TITLE Direct Submission
REFERENCE  3 (bases 1 to 134413)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            TITLE Direct Submission
            JOURNAL Submitted (30-MAY-2003) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
            On Jul 27, 2004 this sequence version replaced gi:45752574.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), Glimmer
            (http://www.tigr.org/tdb/glimmer/glimr_form.html), RiceHMM

```

(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologues of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of B1054C11 clone has an overlap with B1106B03 (DBJ: AP006449) clone at 5' end and an overlap with B1043F11 (DBJ: AP006156) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES

source 1. 141529

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="Genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="9"

/clone="B1054C11"

complement(779..1069)

/gene="B1054C11.1"

complement(779..1069)

/gene="B1054C11.1"

/note="hypothetical ORF predicted by GlimmerM

this category is not included in IRGSP standard"

4054..6208

/gene="B1054C11.2"

4054..6208

/gene="B1054C11.2"

/note="probably inactive due to including frameshift(s) in CDS

probably inactive due to including stop codon(s) in CDS

pseudogene, putative reverse transcriptase"

/pseudo

7971..8195

/gene="B1054C11.3"

<7971..>8195

/gene="B1054C11.3"

/note="start and end point are not identified"

7971..8195

/gene="B1054C11.3"

/note="predicted by GlimmerM etc."

/codon_start=1

/product="hypothetical protein"

/protein_id="BAD34388.1"

/db_xref="GI:50726671"

/translation="MAKLDTLRMQEQQQAYEARQAKHTQQFYAQPVVPYAAQ APYVFAKPAQVAVRLFPQTSEVPRTAP"

8808..10491

/gene="B1054C11.4"

8808..10491

/gene="B1054C11.4"

/note="probably inactive due to 5'exon missing in CDS

probably inactive due to including stop codon(s) in CDS

pseudogene, RIRE2 GAG-POL precursor orf1"

/pseudo

10576..11713

/gene="B1054C11.5"

10576..11713

/gene="B1054C11.5"

/note="probably inactive due to 3'exon missing in CDS

probably inactive due to including frameshift(s) in CDS

probably inactive due to including stop codon(s) in CDS

pseudogene, RIRE2 GAG-POL precursor orf2"

/pseudo

complement(join(12879..12958,13076..13343,17753..17915,18075..18227,18368..18485,18555..18642))

/gene="B1054C11.6"

complement(join(12879..12958,13076..13343,17753..17915,18075..18227,18368..18485,18555..18642))

/gene="B1054C11.6"

/note="hypothetical ORF predicted by GENSCAN

this category is not included in IRGSP standard"

20255..20710

/gene="B1054C11.9"

20255..20710

/gene="B1054C11.9"

/note="5' Terminal repeat"

/rpt_type=Terminal

22332..22823

/gene="B1054C11.7"

22332..22823

/gene="B1054C11.7"

/note="pol polyprotein-like"

24296..28412

/gene="B1054C11.9"

24296..27636

/gene="B1054C11.9"

/gene="B1054C11.9"

/note="probably inactive due to including frameshift(s) in CDS

pseudogene, RIRE2 GAG-POL precursor orf2"

/pseudo

27958..28412

/gene="B1054C11.9"

/note="3' Terminal repeat"

/rpt_type=Terminal

36455..36500)

complement(join(30838..30884,31676..31753,32773..32922,36455..36500))

/gene="B1054C11.10"

complement(join(30838..30884,31676..31753,32773..32922,36455..36500))

/gene="B1054C11.10"

/note="hypothetical ORF predicted by GENSCAN

this category is not included in IRGSP standard"

complement(36818..37609)

/gene="B1054C11.11"

complement(36818..37609)

/gene="B1054C11.11"

/note="probably inactive due to including frameshift(s) in CDS

probably inactive due to including stop codon(s) in CDS

pseudogene, putative polyprotein"

/pseudo

complement(join(39930..39959,40065..40202))

/gene="B1054C11.12"

complement(join(39930..39959,40065..40202))

/gene="B1054C11.12"

/note="hypothetical ORF predicted by GlimmerM

this category is not included in IRGSP standard"

```

gene      join(42044..42502,42615..43076)
misc_feature
/gene="B1054C11.13"
join(42044..42502,42615..43076)
/gene="B1054C11.13"
/notes="transposase-like protein"
43456..43906
/gene="B1054C11.14"
join(43456..43906)
/gene="B1054C11.14"
/notes="start and end point are not identified"
43456..43906
/gene="B1054C11.14"
/notes="predicted by GeneMark.hmm etc."
/codon start=1
/product="hypothetical protein"
/protein.id="BAD34389.1"
/db xref="GI:50726672"
/translation="MLAKAAQKNPKGHTPKKGGSPIEKRRASNTALEKILVEL
LHBNTPYRQGWTSSEAWNKIVKFHEKDRYVCLTKSQIQEKELKREYRMLKEA
RKSGASWNNQRICIEAPAI"
45057..47758
/gene="B1054C11.15"
join(45057..47758)
/gene="B1054C11.15"

Query Match      91.6%  Score 17.4;  DB 4;  Length 141529;
Best Local Similarity 94.7%;  Pred. No. 3.5e+02;
Matches 18;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

QY      1  CTTTACTTCATAGTCCTTG 19
          |||||
Db      58588  CTTTACTTCATAGTCCTCG 58570

RESULT 46
AC091155  AC091155  144430 bp  DNA  linear  PRI 24-APR-2002
LOCUS      Homo sapiens chromosome 17, clone CTD1-2306D4, complete sequence.
DEFINITION AC091155
ACCESSION  AC091155
VERSION    AC091155.5  GI:20128191
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 144430)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,
Camarata,J., Collamore,A., Cooke,P., Dearellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Retta,R.,
Rise,C., Rogov,P., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,H.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

REFERENCE
Submitted (24-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 144430)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Collins,S., Collamore,A., Cooke,P., Dearellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Retta,R.,
Rise,C., Rogov,P., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,H.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

TITLE
JOURNAL
REFERENCE
AUTHORS

```

JOURNAL REFERENCE AUTHORS

Submitted (01-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 144430)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,H., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 144430)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,H., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On April 10, 2002 this sequence version replaced gi:19683224.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13231
Center clone name: 2306_D_4
----- Location/Qualifiers
1. 144430
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"

FEATURES source

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Cataneese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-724M22;
actual end is at base position 150013 of RP11-724M22.

FEATURES

source

```
1. 150013
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-724M22"
/clone_lib="RPCI-11"
repeat_region 1604..1629
/rpt_family="(T)n"
repeat_region 1940..2102
/rpt_family="MER2_type"
repeat_region 3157..3344
/rpt_family="MIR"
repeat_region 4358..4485
/rpt_family="MER2_type"
repeat_region 4691..4928
/rpt_family="MER1_type"
repeat_region 4938..5019
/rpt_family="MIR"
repeat_region 5020..5323
/rpt_family="Alu"
repeat_region 5324..5414
/rpt_family="MIR"
repeat_region 5843..6691
/rpt_family="L2"
repeat_region 6876..6913
/rpt_family="AT_rich"
repeat_region 6949..6996
/rpt_family="(TAGA)n"
repeat_region 7820..8324
/rpt_family="L2"
repeat_region 8321..8405
/rpt_family="MIR"
repeat_region 8600..8628
/rpt_family="(TC)n"
misc_feature 9468..9733
/note="similar to EST BE841757 (NID:g10274135)"
repeat_region 9666..10139
/rpt_family="MaLR"
misc_feature 9780..10467
/note="similar to EST AV731443 (NID:g10840864)"
repeat_region 11206..11282
/rpt_family="Alu"
```

```
11375..11770
/rpt_family="L2"
repeat_region 11789..11956
/rpt_family="MER1_type"
repeat_region 12024..12298
/rpt_family="L2"
repeat_region 12299..12612
/rpt_family="Alu"
repeat_region 12400..12434
/rpt_family="AT_rich"
repeat_region 12582..12622
/rpt_family="(GAAA)n"
repeat_region 12813..12839
/rpt_family="L2"
repeat_region 12885..13093
/rpt_family="MIR"
repeat_region 13626..13755
/rpt_family="MIR"
repeat_region 14039..14306
/rpt_family="Alu"
repeat_region 14307..14716
/rpt_family="ERVL"
repeat_region 16531..16825
/rpt_family="Alu"
repeat_region 16922..17123
/rpt_family="MIR"
repeat_region 18268..18649
/rpt_family="MER2_type"
repeat_region 18650..18762
/rpt_family="Alu"
repeat_region 18763..18980
/rpt_family="MER2_type"
repeat_region 20151..20274
/rpt_family="MIR"
misc_feature 20616..20947
/note="similar to EST AI418542 (NID:g4264473) tg37a05.x1"
misc_feature 21519..21975
/note="similar to EST N70371 (NID:g1226951) z444f11.s1"
repeat_region 21749..22169
/note="similar to EST W05012 (NID:g1277732) za44f11.r1"
repeat_region 21988..22167
/rpt_family="MIR"
repeat_region 22528..22753
/rpt_family="MIR"
repeat_region 23375..23507
/rpt_family="MIR"
repeat_region 23705..23800
/rpt_family="MIR"
repeat_region 23794..23840
/rpt_family="(CA)n"
repeat_region 23843..24138
/rpt_family="Alu"
repeat_region 24148..24176
/rpt_family="AT_rich"
repeat_region 24160..25313
/rpt_family="L1"
repeat_region 26631..26655
/rpt_family="(T)n"
repeat_region 27027..27049
/rpt_family="(A)n"
repeat_region 30559..30724
/rpt_family="MIR"
repeat_region 32770..33273
/rpt_family="MaLR"
repeat_region 33431..33723
/rpt_family="Alu"
repeat_region 33919..34025
/rpt_family="CR1"
repeat_region 34049..34213
/rpt_family="MIR"
repeat_region 34480..34782
/rpt_family="Alu"
repeat_region 34762..34782
```

```

repeat_region /rpt family=" (A)n"
34789. .34936
repeat_region /rpt family="MIR"
35145. .35269
repeat_region /rpt family="MIR"
35521. .35828
repeat_region /rpt family="Alu"
35807. .35828
repeat_region /rpt family=" (A)n"
36448. .36486

```

Query Match 91.6%; Score 17.4; DB 5; Length 150013;

Best Local Similarity 94.7%; Pred. No. 3.5e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 CTTTACTTCATGCTTTG 19
| |||||
Db 97641 CATTACTTCATGCTTTG 97623

```

RESULT 48

AP003856 LOCUS AP003856 158479 bp DNA linear PLN 05-OCT-2004
 DEFINITION Oriza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
 PAC clone:P0697D09.
 ACCESSION AP003856 BA000010
 VERSION AP003856.4 GI:20161655

ORYZA SATIVA (japonica cultivar-group)

Oriza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoidae; Oryzaceae; Oryza.

REFERENCE

1
 AUTHORS Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
 Katayose,Y., Wu,J., Niimura,Y., Cheng,Z., Nagamura,Y.,
 Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
 Chiden,Y., Hayaishi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
 Hanada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y., Ito,Y.,
 Itonuma,A., Iijima,M., Ikeda,M., Ikeno,M., Ito,S., Ito,T., Ito,Y.,
 Ito,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katagiri,S.,
 Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
 Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
 Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
 Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
 Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
 Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
 Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
 Yano,M., Jiang,J. and Gojobori,T.
 The genome sequence and structure of rice chromosome 1
 Nature 420 (6913), 312-316 (2002)

2 (bases 1 to 158479)

SAKAKI,T., MATSUMOTO,T. and YAMAMOTO,K.

Direct Submission

Submitted (05-JUL-2001) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)

On Apr 16, 2002 this sequence version replaced gi:16904699.
 Genes were predicted from the integrated results of the following:

GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH

(http://www.softberry.com/), GeneMark.hmm

(http://opal.biology.gatech.edu/GeneMark/), Glimmer

(http://opal.biology.gatech.edu/glimmer/glmr form.html), RiceHMM

(http://www.tigr.org/tdb/glimmer/glmr form.html), SplicePredictor

(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor

(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4

(http://globin.cse.psu.edu/html/docs/sim4.html), gap2

(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI NonRedundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DDBJ. Protein homologies of the coding

regions were searched against NCBI NonRedundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from SP6 to T7 of the PAC clone.
 This sequence of P0697D09 clone has an overlap with P0503E05 (DDBJ:
 AP003021) clone at 5' end. Detailed information on overlap and
 assembly quality together with annotation of this entry is
 available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES

source

1. 158479

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="1"

/clones="P0697D09"

1754. 1984

/gene="P0697D09.1"

1754. 1984

/gene="P0697D09.1"

/note="hypothetical ORF

predicted by GlimmerM

this category is not included in IRGSP standard"

complement(2282. .2515)

/gene="P0697D09.2"

complement(2282. .2515)

/gene="P0697D09.2"

/note="hypothetical ORF

predicted by GlimmerM

this category is not included in IRGSP standard"

complement(2588. .3167)

/gene="P0697D09.3"

complement(join(<2588. .2913,3032. .>3167))

/note="start and end point are not identified"

/gene="P0697D09.3"

complement(join(2588. .2913,3032. .3167))

/note="predicted by GeneMark.hmm etc."

/codon_start=1

/product="hypothetical protein"

/protein_id="BAD53954.1"

/db_xref="GI:53793491"

/translation="MERGCAETELAAAKPAVEAAHQSGSGGGGAAWLATATER

GALAGKSSPESGARRGESAGLGNKNGDEGSTGVVYMGAGEGDRQGRQGGN

GRDVGATWARGGKPAIPAGRVGEKVEEWATTARSVDVRVGGRRRQF"

complement(3399. .3725)

/gene="P0697D09.4"

complement(<3399. .>3725)

/gene="P0697D09.4"

/note="start and end point are not identified"

complement(3399. .3725)

/gene="P0697D09.4"

/note="predicted by GeneMark.hmm etc."

/codon_start=1

/product="hypothetical protein"

/protein_id="BAD53955.1"

/db_xref="GI:53793492"

/translation="MGVHGRPKYRRGAAANDSGARPQDGHRRHRWRSARGERR

KERRRGGAHQARHQREARRRRSSWHGNGPVPFREREPAADGGKDLGNKEGLGN

REEAHE"


```
gene      join(4678..4782,5254..5361)
misc_feature /genes="P0697D09.5"
            join(4678..4782,5254..5361)
            /genes="P0697D09.5"
            /notes="hypothetical ORF
            predicted by GeneMark.hmm
            this category is not included in IRGSP standard"
            7882..8256
gene      /genes="P0697D09.6"
mRNA     <7882..>8256
         /genes="P0697D09.6"
         /note="start and end point are not identified"
         7882..8256
CDS      /genes="P0697D09.6"
         /notes="predicted by GlimmerM etc."
         /codon_start=1
         /product="hypothetical protein"
         /protein_id="BAD53956.1"
         /db_xref="GI:53793493"
         /translation="MBEGDAGRRRWPWRCSPAHDAEVPAESDGRGVDGVVTRTTAN
         ATRLGAVASGRSGGGGGKGGDGARAMEGTGDTKRRERGGEPLRYPAGVDRGW
         DGRDFAGVGGRGGEEREAGFKN"
         complement(join(8318..12103,17938..18448))
gene      /genes="P0697D09.7"
repeat_unit complement(8318..8835)
         /genes="P0697D09.7"
         /notes="3' terminal repeat"
         /rpt_type="terminal"
         complement(8863..12103)
misc_feature /genes="P0697D09.7"
         /notes="probably inactive due to including stop codon(s) in
         CDS
         pseudogene, polyprotein"
         /pseudogene
         14751..15035
gene      /genes="P0697D09.8"
misc_feature 14751..15035
         /genes="P0697D09.8"
         /notes="hypothetical ORF
         predicted by GlimmerM
         this category is not included in IRGSP standard"
         16300..17385
gene      /genes="P0697D09.9"
mRNA     <16900..>17385
         /genes="P0697D09.9"
         /notes="start and end point are not identified"
         16900..17385
CDS      /genes="P0697D09.9"
         /notes="similar to Oryza sativa chromosome 1, OJ1460_H08.6"
         /codon_start=1
         /product="hypothetical protein"
         /protein_id="BAD53957.1"
         /db_xref="GI:53793494"
         /translation="MAVTRVAVRGSGPGTGGGAPRGTDVVPTWTRGGHVARHAG
         CRTDQAPIPRVDRTRAAKAAAGAHRLKPARPWTALPRGGGRRRPTASGGGA
         SAKAMLAALKWLGCGHGMRYRGAAADGGRRRAAKWLGGVATKRRGRERNGR
         G"
         complement(17938..18448)
repeat_unit /genes="P0697D09.7"
         /notes="5' terminal repeat"
         /rpt_type="terminal"
         join(19302..19350,19637..19840,20866..21059)
gene      /genes="P0697D09.10"
misc_feature join(19302..19350,19637..19840,20866..21059)
         /genes="P0697D09.10"
         /notes="hypothetical ORF
         predicted by GENSCAN
         this category is not included in IRGSP standard"
         complement(24099..25801)
gene      /genes="P0697D09.11"
misc_feature complement(24099..25801)
         /genes="P0697D09.11"
         /note="probably inactive due to 5' exon missing in CDS
```

```
gene      pseudogene, wall-associated kinase 5"
         /pseudogene
         join(27260..27262,29203..29946)
Query Match 91.6%; Score 17.4; DB 4; Length 158479;
Best Local Similarity 94.7%; Pred No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY - 1 CTTTACTTCATAGCTTTTG 19
      |||||
Db 105161 CTTTGCTTCATAGCTTTTG 105179
      |||||
RESULT 49
AP003556/c
LOCUS      AP003556                167007 bp    DNA    linear    HTG 02-MAY-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-644G15 map 11q, WORKING DRAFT
SEQUENCE, 28 unordered pieces.
ACCESSION AP003556
VERSION   AP003556.1 GI:13928405
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 167007)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
          Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (27-APR-2001) Masahira Hattori, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
          ----- Genome Center
          Center: RIKEN Genomic Sciences Center (GSC)
          Center code: RIKEN
          Web site: http://hgp.gsc.riken.go.jp/
          Contact: hattori@gsc.riken.go.jp
          ----- Project Information
          Center project name: HumDraft11
          Center clone name: RP11-644G15
          ----- Summary Statistics
          Sequencing vector: PCR products; 100% of reads
          Chemistry: Dye-terminator ET-amersham; 100% of reads
          Assembly program: Phrap; version 0.990329
          Consensus quality: 161481 bases at least Q40
          Consensus quality: 163250 bases at least Q30
          Consensus quality: 163985 bases at least Q20
          Insert size: 164307; sum-of-contigs
          Quality coverage: 8.64x in Q20 bases; sum-of-contigs
          -----
NOTE: This is a 'working draft' sequence. It currently consists of
28 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 16343 contig of 16343 bp in length
33154 contig of 16711 bp in length
33255 45637 contig of 12383 bp in length
58174 contig of 12437 bp in length
58275 69718 contig of 11444 bp in length
69819 77379 contig of 7189 bp in length
77480 84668 contig of 7189 bp in length
84769 92027 contig of 7259 bp in length
92128 99269 contig of 7142 bp in length
99370 104879 contig of 5510 bp in length
110490 111041 contig of 6062 bp in length
111142 116604 contig of 5463 bp in length
116705 122298 contig of 5594 bp in length
```

122399 126480 contig of 4082 bp in length
126581 130116 contig of 3536 bp in length
130217 133305 contig of 3089 bp in length
133406 138289 contig of 4884 bp in length
138390 141974 contig of 3585 bp in length
142075 145634 contig of 3560 bp in length
145735 148493 contig of 2759 bp in length
148594 151947 contig of 3354 bp in length
152048 155192 contig of 3145 bp in length
155293 157139 contig of 1847 bp in length
157240 159731 contig of 2492 bp in length
159832 162489 contig of 2658 bp in length
162590 164577 contig of 1988 bp in length
164678 166329 contig of 1652 bp in length
166430 167007 contig of 578 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 16343: contig of 16343 bp in length
* 16344 16443: gap of 100 bp
* 16444 33154: contig of 16711 bp in length
* 33155 33254: gap of 100 bp
* 33255 45637: contig of 12383 bp in length
* 45638 45737: gap of 100 bp
* 45738 58174: contig of 12437 bp in length
* 58175 58274: gap of 100 bp
* 58275 69718: contig of 11444 bp in length
* 69719 69818: gap of 100 bp
* 69819 77379: contig of 7561 bp in length
* 77380 77479: gap of 100 bp
* 77480 84668: contig of 7189 bp in length
* 84669 92027: contig of 7259 bp in length
* 84769 92027: gap of 100 bp
* 92028 92127: gap of 100 bp
* 92128 99269: contig of 7142 bp in length
* 99270 99369: gap of 100 bp
* 99370 104879: contig of 5510 bp in length
* 104880 104979: gap of 100 bp
* 104980 111041: contig of 6062 bp in length
* 111042 111141: gap of 100 bp
* 11142 116604: contig of 5463 bp in length
* 116605 116705: gap of 100 bp
* 116705 122298: contig of 5594 bp in length
* 122299 122399: gap of 100 bp
* 122399 126480: contig of 4082 bp in length
* 126481 126581: gap of 100 bp
* 126581 130116: contig of 3536 bp in length
* 130117 130216: gap of 100 bp
* 130217 133305: contig of 3089 bp in length
* 133306 133406: gap of 100 bp
* 133406 138289: contig of 4884 bp in length
* 138290 138390: gap of 100 bp
* 138390 141974: contig of 3585 bp in length
* 141975 142075: gap of 100 bp
* 142075 145634: contig of 3560 bp in length
* 145635 145735: gap of 100 bp
* 145735 148493: contig of 2759 bp in length
* 148494 148594: gap of 100 bp
* 148594 151947: contig of 3354 bp in length
* 151948 152048: gap of 100 bp
* 152048 155192: contig of 3145 bp in length
* 155193 155293: gap of 100 bp
* 155293 157139: contig of 1847 bp in length
* 157140 157240: gap of 100 bp
* 157240 159731: contig of 2492 bp in length
* 159732 159831: gap of 100 bp
* 159832 162489: contig of 2658 bp in length
* 162490 162589: gap of 100 bp

* 162590 164577: contig of 1988 bp in length
* 164578 164678: gap of 100 bp
* 164678 166329: contig of 1652 bp in length
* 166330 166429: gap of 100 bp
* 166430 167007: contig of 578 bp in length.

FEATURES
 source
 1. 167007
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-644G15"
 1. 16343
 /notes="assembly_fragment"
 16444..33154
 /notes="assembly_fragment"
 33255..45637
 /notes="assembly_fragment"
 45738..58174
 /notes="assembly_fragment"
 58275..69718
 /notes="assembly_fragment"
 69819..77379
 /notes="assembly_fragment"
 77480..84668
 /notes="assembly_fragment"
 84769..92027
 /notes="assembly_fragment"
 92128..99269
 /notes="assembly_fragment"
 99370..104879
 /notes="assembly_fragment"
 104980..111041
 /notes="assembly_fragment"
 111142..116604
 /notes="assembly_fragment"
 116705..122298
 /notes="assembly_fragment"
 122399..126480
 /notes="assembly_fragment"
 126581..130116
 /notes="assembly_fragment"
 130217..133305
 /notes="assembly_fragment clone_end:SP6 vector_side:left"
 133406..138289
 /notes="assembly_fragment"
 138390..141974
 /notes="assembly_fragment"
 142075..145634
 /notes="assembly_fragment"
 145735..148493
 /notes="assembly_fragment"
 148594..151947
 /notes="assembly_fragment"
 152048..155192
 /notes="assembly_fragment"
 155293..157139
 /notes="assembly_fragment"
 157240..159731
 /notes="assembly_fragment"
 159832..162489
 /notes="assembly_fragment"
 162590..164577
 /notes="assembly_fragment"
 164678..166329
 /notes="assembly_fragment"
 166430..167007
 /notes="assembly_fragment clone_end:17 vector_side:left"

Query Match 91.6%; Score 17.4; DB 12; Length 167007;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 CTTTACTTCATAGCTTTG 19
Db 23295 CTTTACTTCATGGCTTTG 23277

RESULT 50
AC173916/c
LOCUS
DEFINITION AC173916 167341 bp DNA linear HTG 07-DEC-2005
Aelerix albiventris clone LB4-10706, WORKING DRAFT SEQUENCE, 11
unordered pieces.
ACCESSION AC173916
VERSION AC173916.1 GI:83133777
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Aelerix albiventris (Middle-African hedgehog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Insectivora; Erinaceidae;
Erinaceinae; Aelerix.
1 (bases 1 to 167341)
Antonelias,A., Ayele,K., Bass,D., Benjamin,B., Bera,J.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G.,
Coleman,H., Franks,S., Fuksenko,T., Gestole,M., Greene,A., Guan,X.,
Gupta,J., Gurson,N., Haghighi,P., Han,E., Han,J., Hansen,N.,
Ho,S.-L., Hu,P., Hunter,G., Hurie,B., Idol,J.R., Kwong,P.,
Laric,P., Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M.,
Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B.,
McDowell,J., Montemayor,C., Mullikin,J.C., Park,M., Portnoy,M.E.,
Prasad,A., Puri,O., Rantz,K., Reddix-Dugue,N., Sante,A.,
Schandler,K., Schueler,M.G., Sison,C., Stantripop,S., Taye,A.,
Thomas,J.W., Thomas,P.J., Tsipouri,V., Ung,L., Vogt,J.L.,
Wetherby,K.D., Withers,T.R., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 167341)
Green,E.D.
Direct Submission
Submitted (07-DEC-2005) NTH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehghri.nih.gov
----- Project Information
Center project name: lyf
Center clone name: 107006
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 158069 bases at least Q40
Consensus quality: 160024 bases at least Q30
Consensus quality: 161532 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 166341; sum-of-contigs
Quality coverage: 7.56x in Q20 bases; agarose-fp
Quality coverage: 7.50x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
* 2871: contig of 2871 bp in length
* 2872
* 2971: gap of unknown length
* 2972
* 5156: contig of 2185 bp in length
* 5157
* 5256: gap of unknown length
* 5257
* 8163: contig of 2907 bp in length
* 8164
* 8263: gap of unknown length

```

```

* 8264 12517: contig of 4254 bp in length
* 12518 12617: gap of unknown length
* 12618 19985: contig of 7368 bp in length
* 19986 20085: gap of unknown length
* 20086 25452: contig of 5367 bp in length
* 25453 25552: gap of unknown length
* 25553 31075: contig of 5523 bp in length
* 31076 31175: gap of unknown length
* 31176 55107: contig of 23932 bp in length
* 55108 55207: gap of unknown length
* 55208 77292: contig of 22085 bp in length
* 77293 102159: contig of 24767 bp in length
* 102160 102259: gap of unknown length
* 102260 167341: contig of 65082 bp in length.
* 102260 167341: Location/Qualifiers
FEATURES             source
1..167341
/organism="Aelerix albiventris"
/mol_type="genomic DNA"
/db_xref="taxon:9368"
/clone_lib="LB4-10706"
/clone_lib="LB4"
/notes="BAC resource: http://bacpac.chori.org/"
1..2871
/notes="assembly_fragment"
2872..2971
/estimated_length=unknown
2972..5156
/notes="assembly_fragment"
5157..5256
/estimated_length=unknown
5257..8163
/notes="assembly_fragment"
8164..8263
/estimated_length=unknown
8264..12517
/notes="assembly_fragment"
12518..12617
vector_side:right"
12618..19985
/estimated_length=unknown
19986..20085
/notes="assembly_fragment"
20086..25452
/estimated_length=unknown
25453..25552
/notes="assembly_fragment"
25553..31075
/estimated_length=unknown
31076..31175
/notes="assembly_fragment"
31176..55107
/estimated_length=unknown
55108..55207
/notes="assembly_fragment"
55208..77292
/estimated_length=unknown
77293..77392
/estimated_length=unknown
77393..102159
/notes="assembly_fragment"
102160..102259
/estimated_length=unknown
102260..167341
/notes="assembly_fragment"
vector_side:right"
ORIGIN
Query Match          91.6%; Score 17.4; DB 12; Length 167341;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;

```

Matches	18;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Oy	1	CTTTACTTCATAGTCTTTG	19						
Db	144700	CTTTACTTCATAGTCTTTG	144682						
RESULT 51									
LOCUS	AP003021								
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0503E05.								
ACCESSION	AP003021	BA000010							
VERSION	AP003021.2	GI:13486765							
KEYWORDS									
SOURCE	Oryza sativa (japonica cultivar-group)								
ORGANISM	Oryza sativa (japonica cultivar-group)								
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzaceae; Oryza.								
AUTHORS	1 Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katsuyoshi, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Antonito, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hanada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, S., Ito, T., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yanagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y., Yano, M., Jiang, J. and Gojobori, T.								
TITLE	The genome sequence and structure of rice chromosome 1								
JOURNAL	Nature 420 (6913), 312-316 (2002)								
PUBLISHED	12447438								
REFERENCE	2 (bases 1 to 168258)								
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.								
TITLE	Direct Submission								
JOURNAL	Submitted (06-DEC-2000) Takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan								
COMMENT	(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468) On Mar 28, 2001 this sequence version replaced gi:11602828. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark (http://opal.biology.gatech.edu/GeneMark/), GlimmerH (http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplinePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequences were searched at RGP or DDBJ. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative', and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene								

FEATURES	Location/Qualifiers	source
1..168258	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="genomic DNA"	
	/cultivar="Nipponbare"	
	/db_xref="taxon:39947"	
	/chromosome="1"	
gene	/clone="P0503E05"	
	join(1659..1716,1912..2030)	
	/gene="P0503E05.1"	
misc_feature	join(1659..1716,1912..2030)	
	/gene="P0503E05.1"	
	/note="hypothetical ORF predicted by GENSCAN	
gene	this category is not included in IRGSP standard"	
	complement(2316..4423)	
	/gene="P0503E05.2"	
misc_feature	complement(2316..4423)	
	/gene="P0503E05.2"	
	/note="contains full-length cDNA(s): AK071457 non-coding transcript probably inactive due to including stop codon(s) in CDS"	
gene	complement(4443..5185)	
	/gene="P0503E05.3"	
mrna	complement(4443..5185)	
	/gene="P0503E05.3"	
	/note="supported by full-length cDNA(s): AK107733"	
CDS	complement(4541..4924)	
	/gene="P0503E05.3"	
	/note="contains full-length cDNA(s): AK107733"	
	/codon_start=1	
	/product="unknown protein"	
	/protein_id="BAD52626.1"	
	/db_xref="GI:53791504"	
	/translation="MSDSLPRPAEKRRRRSPRLLPMWCCACVRRRRRCKKHAGAP ASAPALVPASAPAPASRCRCRCPKAPACAVEAGGVGADGLLERAPHRDSRGG RGEREREYDRWHVEGILYRQKLK"	
gene	join(5101..5174,5333..5411,5604..5702)	
mrna	/gene="P0503E05.4"	
	join(<5101..5174,5333..5411,5604..5702)	
	/gene="P0503E05.4"	
	/note="start and end point are not identified"	
CDS	join(5101..5174,5333..5411,5604..5702)	
	/gene="P0503E05.4"	
	/note="predicted by FGENESH etc."	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="BAD52627.1"	
	/db_xref="GI:53791505"	
	/translation="MGDGGASLLEPKRLDIDSINRENREVRGSMKYEAQVAFGLLF FSKQPFWDINDAOFQKYTVDFVQPLHENDVGVTSRVVF"	
gene	7045..7404	
	/gene="P0503E05.5"	
mrna	<7045..>7404	
	/gene="P0503E05.5"	
	/note="start and end point are not identified"	
CDS	7045..7404	
	/gene="P0503E05.5"	
	/note="predicted by FGENESH etc."	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="BAD52628.1"	
	/db_xref="GI:53791506"	
	/translation="MARKDBAEGLKARLSLPKSKRCGDKDGGGGRKRKKTLYNLVGV	

predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.
The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0503E05 clone has an overlap with P0460H02 (DDBJ: AP003257) clone at 5' end and with P0697D09 (DDBJ: AP003856) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

```
CCSKEGVEVLMLPAPAMERLVAEVARGLVGVDAIHARDRWGGRREHMYRGALIGGP
IILEIFSTFKMPHQAA"
/gene .8260
/misc_feature 8012. .8260
/gene="P0503E05.6"
/gene="P0503E05.6"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
8660. .9384
/gene="P0503E05.7"
/mRNA <8660. .9384
/gene="P0503E05.7"
/misc_feature /notes="supported by full-length cDNA(s) : AK064749"
8660. .9384
/gene="P0503E05.7"
/notes="contains full-length cDNA(s) : AK064749
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
11036. .11562
/gene="P0503E05.8"
/mRNA join(<11036. .11237,11531. .>11562)
/gene="P0503E05.8"
/notes="start and end point are not identified"
11036. .11237,11531. .11562
/gene="P0503E05.8"
/notes="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB52629.1"
/db_xref="GI:53791507"
/translations="MDATLTSVHRQRPSCPHAPKLPSLTSSAQTSSTSSPPGRRRR
MSALVKGQPDPTPSLIAGVIKGIPIHFVHAT"
12016. .15977
/gene="P0503E05.9"
/mRNA join(12016. .12506,15527. .15967)
/gene="P0503E05.9"
/notes="supported by full-length cDNA(s) : AK060356"
12074. .12506,15527. .15977)
/gene="P0503E05.9"
/notes="supported by full-length cDNA(s) : AK059622"
12190. .12506,15527. .15957)
/gene="P0503E05.9"
/notes="contains EST(s) : D47200(S12386), AU096192(S12386)
contains full-length cDNA(s) : AK060356, AK059622
similar to Arabidopsis thaliana, AAM67044"
/codon_start=1
/product="unknown protein"
/protein_id="BAB40002.1"
/db_xref="GI:13486769"
/translations="MESSRGKPGINGSGGAAAFDYSSRRGYTGTGAGALPPLAAGSR
APVPDCCVLRVFLVLTASAVVMAADRSQTTVQIAAGEELAPLRVPTAKWTYS
SAFYVFVANAMVFAFGAALAAVRRRSADVPMVMDGLVAMALLFSVAGAAOFGLLG
ERGNHVRWAKVCDVYGFPCERAAVVALIAFADLVLLMLTILTIHKASSY"
17745. .20869
/gene="P0503E05.10-2"
/misc_feature 17745. .20869
/gene="P0503E05.10-2"
/notes="contains full-length cDNA(s) : AK071463
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
17745. .20846
/gene="P0503E05.10-3"
Query Match 91.6%; Score 17.4; DB 4; Length 168258;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGCTTTG 19
||||| |||||||
Db 156379 CTTTGCCTCATAGCTTTG 156397
```

```
RESULT 52
AC015570
LOCUS
DEFINITION
AC015570.5 GI:7329242
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 168509)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-453F17
Unpublished
2 (bases 1 to 168509)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castie,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Perreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 26, 2000 this sequence version replaced gi:6778593.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1024
Center clone name: 453 F.17
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161651 bases at least Q40
Consensus quality: 164748 bases at least Q30
Consensus quality: 166260 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 167609; sum-of-contigs
Quality coverage: 5.2 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-contigs
-----
```

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4392: contig of 4392 bp in length
* 4393 4492: gap of 100 bp
* 4493 8988: contig of 4496 bp in length
* 8989 9088: gap of 100 bp
* 9089 18712: contig of 9624 bp in length
* 18713 18812: gap of 100 bp
```



```

* 3882 3981: gap of 100 bp
* 3982 6990: contig of 3009 bp in length
* 6991 7090: gap of 100 bp
* 7091 11767: contig of 4677 bp in length
* 11768 11867: gap of 100 bp
* 11868 16650: contig of 4783 bp in length
* 16651 16750: gap of 100 bp
* 16751 20824: contig of 4074 bp in length
* 20825 20924: gap of 100 bp
* 20925 26002: contig of 5078 bp in length
* 26003 26102: gap of 100 bp
* 26103 33403: contig of 7301 bp in length
* 33404 33503: gap of 100 bp
* 33504 40903: contig of 7400 bp in length
* 40904 41003: gap of 100 bp
* 41004 51482: contig of 10479 bp in length
* 51483 51582: gap of 100 bp
* 51583 66690: contig of 15108 bp in length
* 66691 66790: gap of 100 bp
* 66791 78793: contig of 12003 bp in length
* 78794 78893: gap of 100 bp
* 78894 97796: contig of 18903 bp in length
* 97797 97896: gap of 100 bp
* 97897 125632: contig of 27736 bp in length
* 97898 125633: gap of 100 bp
* 125634 125733: contig of 42971 bp in length.
* 125733 168703: contig of 42971 bp in length.
FEATURES
    source
        1..168703
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11"
            /clone="RP11-136114"
            /clone.lib="RPC1-11 Human Male BAC"
        1..3881
            /note="assembly_fragment"
        3882..3981
            /estimated_length=100
        3982..6990
            /note="assembly_fragment"
        6991..7090
            /estimated_length=100
        7091..11767
            /note="assembly_fragment"
        11768..11867
            /estimated_length=100
        11868..16650
            /note="assembly_fragment"
        16651..16750
            /estimated_length=100
        16751..20824
            /note="assembly_fragment"
        20825..20924
            /estimated_length=100
        20925..26002
            /note="assembly_fragment"
        26003..26102
            /estimated_length=100
        26103..33403
            /note="assembly_fragment"
        33404..33503
            /estimated_length=100
        33504..40903
            /note="assembly_fragment"
        40904..41003
            /estimated_length=100
        41004..51482
            /note="assembly_fragment"
        51483..51582
            /estimated_length=100
        51583..66690
            /note="assembly_fragment"

```

```

gap 66691..56790
    /estimated_length=100
misc_feature 66791..78793
    /note="assembly_fragment"
gap 78794..78893
    /estimated_length=100
misc_feature 78894..97796
    /note="assembly_fragment"
gap 97797..97896
    /estimated_length=100
misc_feature 97897..125632
    /note="assembly_fragment"
    clone_end:SP6
    vector_side:right"
gap 125633..125732
    /estimated_length=100
misc_feature 125733..168703
    /note="assembly_fragment"
    clone_end:I7
    vector_side:right"

ORIGIN
Query Match 91.6%; Score 17.4; DB 12; Length 168703;
Best Local Similarity 94.7%; Pred.No.3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
    |||||
Db 86430 CTTTACTTCATGGCTTTG 86412

RESULT 54
AC124725/c 168983 bp DNA linear ROD 11-NOV-2003
LOCUS Mus musculus BAC clone RP23-459F16 from chromosome 2, complete
DEFINITION
AC124725 AC124725.3 GI:25189024
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE
    1 (bases 1 to 168983)
    Harkins,R., Cotton,M., Mangiapanello,L. and Dignan,G.
    The sequence of Mus musculus BAC clone RP23-459F16
    Unpublished (2001)
REFERENCE
    2 (bases 1 to 168983)
    Wilson,R.
    Sequencing of Mus musculus
    Unpublished (2001)
REFERENCE
    3 (bases 1 to 168983)
    McPherson,J.D. and Waterston,R.H.
    Direct Submission
    Submitted (15-JUN-2002) Genome Sequencing Center, 4444 Forest Park
    Parkway, St. Louis, MO 63108, USA
REFERENCE
    4 (bases 1 to 168983)
    McPherson,J.D. and Waterston,R.H.
    Direct Submission
    Submitted (30-OCT-2002) Genome Sequencing Center, 4444 Forest Park
    Parkway, St. Louis, MO 63108, USA
REFERENCE
    5 (bases 1 to 168983)
    McPherson,J.D. and Waterston,R.H.
    Direct Submission
    Submitted (23-NOV-2002) Genome Sequencing Center, 4444 Forest Park
    Parkway, St. Louis, MO 63108, USA
REFERENCE
    6 (bases 1 to 168983)
    Wilson,R.
    Direct Submission
    Submitted (11-NOV-2003) Department of Genetics, Washington
    University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
    On Nov 23, 2002 this sequence version replaced gi:24421750.
COMMENT

```



```

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics
-----
Center project name: M_BA0459F16
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseawa and Minako Tatenio in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source	Location/Qualifiers
1..168983	
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="2"
	/map="2"
	/clone="RP23-459F16"
	/clone_lib="RPCI-23"
524..652	
	/rpt_family="Alu"
663..733	
	/rpt_family="ID"
755..943	
	/rpt_family="B2"
2009..2063	
	/rpt_family="ERV1"
2755..2908	
	/rpt_family="L1"
2908..3107	
	/rpt_family="L1"
4208..4300	
	/rpt_family="MIR"
4414..4503	
	/rpt_family="Alu"
4779..5170	
	/rpt_family="L2"
5538..5765	
	/rpt_family="ERVK"
6198..7038	
	/rpt_family="L1"
8020..8337	
	/rpt_family="MaLR"
8346..8444	
	/rpt_family="B2"
repeat_region	11192..11327
	/rpt_family="L1"
repeat_region	11330..11636
	/rpt_family="L1"
repeat_region	11706..11776
	/rpt_family="Alu"
repeat_region	11788..11861
	/rpt_family="L1"
repeat_region	11957..12216
	/rpt_family="B4"
repeat_region	14630..15279
	/rpt_family="L1"
repeat_region	15503..15853
	/rpt_family="ERVK"
repeat_region	17183..17239
	/rpt_family="ERVK"
repeat_region	17259..17425
	/rpt_family="B4"
repeat_region	17512..17626
	/rpt_family="Alu"
repeat_region	17673..17841
	/rpt_family="B2"
repeat_region	17988..18189
	/rpt_family="B2"
repeat_region	18398..19686
	/rpt_family="L1"
repeat_region	19680..19758
	/rpt_family="L1"
repeat_region	19963..20043
	/rpt_family="L1"
repeat_region	20098..20334
	/rpt_family="B2"
repeat_region	20358..20734
	/rpt_family="MaLR"
repeat_region	20735..21572
	/rpt_family="MaLR"
repeat_region	23136..23331
	/rpt_family="B2"
repeat_region	23386..23889
	/rpt_family="L1"
repeat_region	24016..24482
	/rpt_family="L1"
repeat_region	24619..24939
	/rpt_family="MaLR"
repeat_region	26223..26268
	/rpt_family="ERV1"
repeat_region	26784..26877
	/rpt_family="L1"
repeat_region	26906..28545
	/rpt_family="L1"
repeat_region	28552..28739
	/rpt_family="L1"
repeat_region	29793..29923
	/rpt_family="B4"
repeat_region	34162..34240
	/rpt_family="MIR"
repeat_region	34838..35122
	/rpt_family="MaLR"
repeat_region	35215..35609
	/rpt_family="MaLR"
repeat_region	36013..36200
	/rpt_family="MaLR"
repeat_region	38402..38629
	/rpt_family="L1"
repeat_region	38632..38884
	/rpt_family="MaLR"
repeat_region	38885..40759
	/rpt_family="L1"
repeat_region	40760..40923
	/rpt_family="B2"
repeat_region	40986..41188
	/rpt_family="L1"
repeat_region	41189..41226

```

/rpt_family="MaLR"
42017. .42125
/rpt_family="Alu"
42163. .42340
/rpt_family="L1"
42799. .43002
/rpt_family="B2"
43338. .43712
/rpt_family="MaLR"
43715. .44006
/rpt_family="MaLR"
43999. .44218
/rpt_family="MaLR"
44257. .44594

Query Match          91.6%; Score 17.4; DB 6; Length 169983;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|| ||||| ||||| |||||
Db 136616 CTCTACTTCATAGTCTTTG 136598

RESULT 55
AP003174
LOCUS          170156 bp DNA linear PRI 15-MAR-2003
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-136I14,
complete sequence.
ACCESSION AP003174
VERSION AP003174.2 GI:17425259
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2001)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (09-FEB-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Dec 7, 2001 this sequence version replaced gi:12862394.
COMMENT Location/Qualifiers
FEATURES
source
1..170156
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-136I14"

ORIGIN

Query Match          91.6%; Score 17.4; DB 5; Length 170156;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|| ||||| ||||| |||||
Db 104965 CTTTACTTCATAGTCTTTG 104983

RESULT 56
AC127007
LOCUS          170956 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-342K24, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
ACCESSION AC127007
VERSION AC127007.3 GI:25139714
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 170956)
AUTHORS Muzny,D., Marie., Metzker,M., Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Cantar,R., Garcia,A., Garner,T., Garza,M.,
Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevata,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,D., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,S., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 170956)
Worley,K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 170956)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department

```

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:23196239.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: RAFO
 Center clone name: CH230-342K24

----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 158599 bases at least Q40
 Consensus quality: 160475 bases at least Q30
 Consensus quality: 161407 bases at least Q20
 Estimated insert size: 161556; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 165449: contig of 165449 bp in length
 * 165450 165549: gap of unknown length
 * 165550 170956: contig of 5407 bp in length.

FEATURES

source	Location/Qualifiers
	1. .170956
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-342K24"
misc_feature	1. .1394
	/note="wgs_end_extension"
	clone_end:T7"
misc_feature	4483. .5354
	/note="clone_boundary"
	clone_end:T7"
	site:
	end_sequence:BZ209789"
misc_feature	72516. .94000
	/note="clone_boundary"
	clone_end:Sp6
	site:
	end_sequence:BZ209791"
misc_feature	163069. .165449
	/note="wgs_end_extension"
	clone_end:Sp6"
gap	165450. .165549
	/estimated_length=unknown
ORIGIN	
	Direct Submission
	Submitted (10-AUG-2005) Broad Institute of MIT and Harvard, 320
	Charles Street, Cambridge, MA 02141, USA
	On Aug 10, 2005 this sequence version replaced gi:67764029.
	Direct Submission
	Submitted (10-AUG-2005) Broad Institute of MIT and Harvard, 320
	Charles Street, Cambridge, MA 02141, USA
	On Aug 10, 2005 this sequence version replaced gi:67764029.

QY 1 CTTTACTTCATAGCTTTG 19
 Db 144157 CTTTATTTCATAGCTTTG 144175

RESULT 57
 AC164008/c
 LOCUS AC164008 171582 bp DNA linear ROD 10-AUG-2005
 DEFINITION Mus musculus chromosome 1, clone RP23-311L17, complete sequence.
 ACCESSION AC164008
 VERSION AC164008.2 GI:72096358
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 171582)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus chromosome 1, clone RP23-311L17
 Unpublished
 2 (bases 1 to 171582)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., MacDonald,P., Marjor,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

Submitted (15-JUN-2005) Broad Institute of MIT and Harvard, 320
 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 171582)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., MacDonald,P., Marjor,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 TITLE
 JOURNAL
 COMMENT

.Query Match 91.6%; Score 17.4; DB 12; Length 170956;
 Best Local Similarity 94.7%; Pred. No. 3.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: J5706
Center clone name: 311_L_17

repeat_region	23962. .24156 /rpt_family="Lx6"
repeat_region	complement(25203. .25355) /rpt_family="RMERSA"
repeat_region	26206. .26246 /rpt_family=" (TTTG)" n
repeat_region	26247. .26467 /rpt_family="L1_Mus3"
repeat_region	complement(26588. .26799) /rpt_family="B4A"
repeat_region	27201. .27246 /rpt_family="AT_rich"
repeat_region	27278. .27334 /rpt_family=" (TTTG)" n
repeat_region	complement(27335. .27481) /rpt_family="B1_Mus2"
repeat_region	complement(27571. .28185) /rpt_family="L1_Mm"
repeat_region	28173. .28610 /rpt_family="Lx7"
repeat_region	28611. .28784 /rpt_family="B2_Mm2"
repeat_region	28785. .28940 /rpt_family="Lx7"
repeat_region	complement(28942. .29231) /rpt_family="RMER4B"
repeat_region	29232. .29336 /rpt_family=" (TAGA)" n
repeat_region	complement(29337. .29408) /rpt_family="RMER4B"
repeat_region	29450. .29802 /rpt_family="Lx7"
repeat_region	29803. .30125 /rpt_family="ORR1A3"
repeat_region	30135. .30355 /rpt_family="Lx6"
repeat_region	30507. .30972 /rpt_family=" (TATG)" n
repeat_region	31000. .31025 /rpt_family=" (TA)" n
repeat_region	31087. .31358 /rpt_family="RMER6C"
repeat_region	31359. .31734 /rpt_family="Lx2"
repeat_region	31735. .32124 /rpt_family="RMER6C"
repeat_region	complement(32375. .32557) /rpt_family="Lx8"
repeat_region	complement(32827. .33407) /rpt_family="Lx9"
repeat_region	complement(33411. .33843) /rpt_family="Lx9"
repeat_region	33894. .33917 /rpt_family=" (TA)" n
repeat_region	33918. .34227 /rpt_family="L1_Rod"
repeat_region	35066. .35096 /rpt_family="AT_rich"
repeat_region	35108. .35176 /rpt_family="Lx9"
repeat_region	35334. .35506 /rpt_family="MER"

Query Match 91.6%; Score 17.4; DB 6; Length 171582;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18: Conservative 0; Mismatches 1; Indels 0;

Qy	1	C	T	T	A	C	T	C	A	T	A	G	T	C	T	T	T	19
Db	158637	C	T	T	T	C	T	C	A	T	A	G	T	C	T	T	158619	

RESULT 58

AC148231/c
 LOCUS
 DEFINITION
 AC148231 176362 bp DNA linear HTG 17-FEB-2004
 ordered pieces.
 AC148231
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AC148231.1 GI:42558291
 HTG; HTGS PHASE2; HTGS DRAFT.
 Didelphis virginiana (North American opossum)
 Didelphis virginiana
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 1 (Bases 1 to 176362)
 Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 Direct Submission
 Unpublished
 2 (Bases 1 to 176362)
 Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 Direct Submission
 Submitted (17-FEB-2004) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: O148
 Bac Clone Name: LB3-44H21

This sequence has been compared to sequences of other species
 using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
 viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=ntvalue-LIPC

The order-orientation of the draft sequence was accomplished by
 using:
 Avid (<http://baboon.math.berkeley.edu/mavid>),
 Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHUBI)

If the Bac Library Name is LB1 to LB4, please see website
 for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>
 These libraries are available through the BACPAC Resources Center:
<http://www.chori.org/bacpac/libraryres.htm> as LBNL-1 to LBNL-4.

Summary Statistics:

Sequencing vector: plasmid; pUC18
 Chemistry: Dye-terminator Big Dye
 Assembly program: Phrap version 0.990329.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 1086: contig of 1086 bp in length
 * 1087 1186: gap of unknown length
 * 1187 17979: contig of 16793 bp in length
 * 17980 18079: gap of unknown length
 * 18080 60719: contig of 42640 bp in length
 * 60820 60819: gap of unknown length
 * 60820 73619: contig of 12800 bp in length
 * 73620 73719: gap of unknown length
 * 73720 87458: contig of 13739 bp in length
 * 87459 87558: gap of unknown length
 * 87559 166569: contig of 79011 bp in length
 * 166570 166669: gap of unknown length
 * 166670 176362: contig of 9693 bp in length.

FEATURES

Location/Qualifiers

source
 1. 176362
 /organism="Didelphis virginiana"
 /mol_type="genomic DNA"
 /db_xref="taxon:9267"
 /clone="LB3-44H21"
 1087..1186
 /estimated_length=unknown
 17980..18079
 /estimated_length=unknown
 60720..60819
 /estimated_length=unknown
 73620..73719
 /estimated_length=unknown
 87459..87558
 /estimated_length=unknown
 166570..166669
 /estimated_length=unknown
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 12; Length 176362;
 Best Local Similarity 94.7%; Pred. No. 3.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 CTTTACTTCATAGTCTTTG 19
 |||||
 Db 71363 CTTTACTTCATAGACTTTG 71345
 RESULT 59
 AL845354
 LOCUS
 DEFINITION
 AL845354 178468 bp DNA linear ROD 24-JUN-2003
 Mouse DNA sequence from clone RP23-318L10 on chromosome 2, complete
 sequence.
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AL845354 GI:32187955
 HTG.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 Tracey, A.
 Direct Submission
 Submitted (24-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Jun 24, 2003 this sequence version replaced GI:32131063.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:
 En:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-318L10 is
 from the RPII-23 Mouse BAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6.

FEATURES
 source
 1. .178468
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-318L10"
 /clone_lib="RPII-23"

ORIGIN

Query Match 91.6%; Score 17.4; DB 6; Length 178468;
 Best Local Similarity 94.7%; Pred. No. 3.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATGCTTTG 19

Db 138089 CTTTACTTCATGCTTTG 138107

RESULT 60

AL512635/c

LOCUS
 DEFINITION

AL512635 178631 bp DNA linear PRI 18-MAY-2005
 Human DNA sequence from clone RPII-15P13 on chromosome 9 Contains a
 survival of motor neuron 1, telomeric (SMN1) pseudogene and the 3'
 end of the MLT3 gene for myeloid/lymphoid or mixed-lineage
 leukemia (trithorax homolog, Drosophila); translocated to, 3 (AF9),
 complete sequence.

ACCESSION

VERSION AL512635.8 GI:16409123

KEYWORDS

HTG; AF9; MLT3; SMN1.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.
 1 (bases 1 to 178631)
 Pearce,A.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK; E-mail enquiries: vegas@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 On Oct 24, 2001 this sequence version replaced gi:15131291.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 En:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>
 RPII-15P13 is from the library RPII-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: vegas@sanger.ac.uk

Draft Sequence Produced by Whitehead Institute/MIT Center for

Genome Research, 320 Charles Street,

Cambridge, MA 02141, USA

<http://www-seq.wi.mit.edu>

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone; and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES

Location/Qualifiers
 1. .178631

source

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="9"

/clone="RPII-15P13"

/clone_lib="RPII-11.1"

misc_feature

1
 /note="Clone_left_end: RPII-15P13"

misc_feature

53156
 /note="Clone_right_end: RPII-594J14"

gene

148943..149774

/locus_tag="RPII-15P13.1-001"

pseudocds

148943..149774

/locus_tag="RPII-15P13.1-001"

/note="match: proteins: AAC83178 AAH15308 O02771 O35876

P97801 Q16637 Q96J51"

pseudocodon_start=1

149506

misc_feature

/locus_tag="RPII-15P13.1-001"

/note="Clone_left_end: RPII-336O12"

gene

join(Complement(AL354879.12:14212..14510),

complement(AL354879.12:12621..12801),

complement(AL513498.2:124696..124778),

complement(AL513498.2:116115..116258),

complement(AL513498.2:81713..82417),

complement(AL513498.2:33661..33736),

complement(AL513498.2:31468..31597),

complement(178239..178338),complement(172305..172376),

complement(171022..171093),complement(159162..164071))

/gene="MLLT3"

mRNA

/locus_tag="RPII-15P13.2-001"

join(Complement(AL354879.12:14212..14510),

complement(AL354879.12:12621..12801),

complement(AL513498.2:124696..124778),

complement(AL513498.2:116115..116258),

complement(AL513498.2:81713..82417),

complement(AL513498.2:33661..33736),

complement(AL513498.2:31468..31597),

complement(178239..178338),complement(172305..172376),

complement(171022..171093),complement(159162..164071))

/gene="MLLT3"

/locus_tag="RPII-15P13.2-001"

/product="myeloid/lymphoid or mixed-lineage leukemia

(trithorax homolog, Drosophila)"; translocated to, 3"

/note="match: ESTs: AA0406302.1 AA42076.1 AA443284.1

AI336478.1 AI554049.1 AI620408.1 AI739108.1 AI768397.1

AL535503.1 AL565947.1 AW044186.1 AW163788.1 AW302520.1

AW451255.1 AW572468.1 BF104863.1 BF510928.1 BG940265.1

BG958893.1 BI550359.1 BI561777.1 BI710808.1 BM018568.1

BU623899.1 CA418678.1 HI8503.1 H93306.1 N56884.1

match: cDNAs: BC036089.1 DI6688.1 LI3744.1"

complement(159162)

polyA_site

/gene="MLLT3"

/locus_tag="RPII-15P13.2-001"

complement(159188..159193)

polyA_signal

/gene="MLLT3"

join(Complement(AL513498.2:50373..50479),

complement(AL513498.2:33661..33736),

complement(AL513498.2:31468..31597),

complement(178239..178338),complement(172305..172376),

complement(162574. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-006"
 join(complement(ALS13498.2:50373. .50479),
 complement(ALS13498.2:33661. .33736),
 complement(ALS13498.2:31468. .31597),
 complement(178239. .178338), complement(172305. .172376),
 complement(162574. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-006"
 /product="myeloid/lymphoid or mixed-lineage leukemia
 (trithorax homolog, Drosophila)"; translocated to, 3"
 /note="match: cDNAs: AK019458.1"
 join(complement(ALS13498.2:31468. .31561),
 complement(178239. .178338), complement(163790. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-007"
 join(complement(ALS13498.2:31468. .31561),
 complement(178239. .178338), complement(163790. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-007"
 /product="myeloid/lymphoid or mixed-lineage leukemia
 (trithorax homolog, Drosophila)"; translocated to, 3"
 /note="match: ESTs: N66033.1"
 join(complement(ALS13498.2:50373. .50463),
 complement(172305. .172376), complement(171022. .171093),
 complement(163850. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-005"
 join(complement(ALS13498.2:50373. .50463),
 complement(172305. .172376), complement(171022. .171093),
 complement(163850. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-005"
 /product="myeloid/lymphoid or mixed-lineage leukemia
 (trithorax homolog, Drosophila)"; translocated to, 3"
 /note="match: ESTs: BU664780.1"
 join(complement(ALS13498.2:50373. .50461),
 complement(ALS13498.2:33661. .33736),
 complement(ALS13498.2:31468. .31597),
 complement(178239. .178338), complement(172305. .172376),
 complement(171022. .171093), complement(163903. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-002"
 join(complement(ALS13498.2:50373. .50461),
 complement(ALS13498.2:33661. .33736),
 complement(ALS13498.2:31468. .31597),
 complement(178239. .178338), complement(172305. .172376),
 complement(171022. .171093), complement(163903. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-002"
 /product="myeloid/lymphoid or mixed-lineage leukemia
 (trithorax homolog, Drosophila)"; translocated to, 3"
 /note="match: ESTs: BF243978.1 BG940264.1 BM793722.1"
 join(complement(ALS13498.2:31468. .31580),
 complement(178239. .178338), complement(172305. .172376),
 complement(163940. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-006"
 /standard name="OTTHUMP0000021129"
 /codon start=1
 /product="myeloid/lymphoid or mixed-lineage leukemia
 (trithorax homolog, Drosophila)"; translocated to, 3"
 /protein_id="CAH70707.1"
 /db_xref="GI:55665053"
 /db_xref="UniProtKB/TrEMBL:Q5VV07"
 /translations="MKDLHSDNEEDSEVEDNDSEMERPNRGGSRRLVSLSDG
 SDESSASPLHPPPLTKTNNQILEVKSPIKQSKSDKIKNGECDKVNLIEE
 TGHFHTINTFFDLGSLDKITVRKQSLVETSGTS"
 join(complement(AL354879.12:14212. .14223),
 join(complement(AL354879.12:12621. .12801),
 complement(ALS13498.2:124696. .124778),
 complement(ALS13498.2:116115. .116258),

complement(ALS13498.2:81713. .82417),
 complement(ALS13498.2:33661. .33736),
 complement(ALS13498.2:31468. .31597),
 complement(178239. .178338), complement(172305. .172376),
 complement(171022. .171093), complement(163940. .164071))
 /gene="MLLT3"
 /locus tag="RP11-15P13.2-002"
 /standard name="OTTHUMP0000021127"
 /note="match: proteins: BAA04090 P42568 Q14768 Q8VDR6
 Q99MK4"
 Query Match 91.6%; Score 17.4; DB 5; Length 178631;
 Best Local Similarity 94.7%; Pred. No. 3.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGCTTTG 19
 Db 66553 CTTTACTTCATCGCTTTG 66535
 RESULT 61
 AC011949
 LOCUS
 DEFINITION Homo sapiens clone RP11-17E12, WORKING DRAFT SEQUENCE, 10 unordered
 pieces.
 AC011949
 AC011949.3 GI:7387331
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 VERSION
 KEYWORDS Homo sapiens
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 179112)
 Birren, B., Linton, L., Nussbaum, C. and Lander, E.
 Homo sapiens chromosome, clone RP11-17E12
 Unpublished
 2 (bases 1 to 179112)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhaider, B.,
 Brown, A., Castelle, A., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-OCT-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 3, 2000 this sequence version replaced gi.6453950.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L3628
 Center clone name: 17 E 12
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 150637 bases at least Q40
 Consensus quality: 166570 bases at least Q30

Consensus quality: 173803 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 178212; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 3511: contig of 3511 bp in length
* 3512: gap of 100 bp
* 3611: gap of 100 bp
* 3612 6703: contig of 3092 bp in length
* 6704 6803: gap of 100 bp
* 6804 12436: contig of 5633 bp in length
* 12437 12536: gap of 100 bp
* 12537 18649: contig of 6113 bp in length
* 18650 18749: gap of 100 bp
* 18750 31179: contig of 12430 bp in length
* 31180 31279: gap of 100 bp
* 31280 44707: contig of 13428 bp in length
* 44708 4807: gap of 100 bp
* 4808 69716: contig of 24909 bp in length
* 69717 69816: gap of 100 bp
* 69817 99053: contig of 29236 bp in length
* 99054 99153: gap of 100 bp
* 99154 132097: contig of 32945 bp in length
* 132098 132197: gap of 100 bp
* 132198 179112: contig of 46915 bp in length.

FEATURES

source
1. 179112
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-17E12"
/clone_lib="RPCI-11 Human Male BAC"
misc_feature
1. 3511
/note="assembly_fragment"
gap
3512..3611
/estimated_length=100
misc_feature
3612..6703
/note="assembly_fragment"
gap
6704..6803
/estimated_length=100
misc_feature
6804..12436
/note="assembly_fragment"
gap
12437..12536
/estimated_length=100
misc_feature
12537..18649
/note="assembly_fragment"
gap
18650..18749
/estimated_length=100
misc_feature
18750..31179
/note="assembly_fragment"
gap
31180..31279
/estimated_length=100
misc_feature
31280..44707
/note="assembly_fragment"
gap
44708..4807
/estimated_length=100
misc_feature
4808..69716
/note="assembly_fragment"
clone_end:T7
vector_side:left
69717..69816
/estimated_length=100
misc_feature
69817..99052
/note="assembly_fragment"
clone_end:SP6

vector_side:right"
99053..99152
/estimated_length=100
misc_feature
99153..132097
/note="assembly_fragment"
gap
132098..132197
/estimated_length=100
misc_feature
132198..179112
/note="assembly_fragment"

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 179112;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 42157 CTTTACTTCATAGTCTTGG 42175

RESULT 62

AC022743/c 182466 bp DNA linear PRI 01-MAY-2002
LOCUS Homo sapiens chromosome 18, clone RP11-789C7, complete sequence.
DEFINITION

AC022743
AC022743.7 GI:20148189
VERSION HTG.
KEYWORDS

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE

1 (bases 1 to 182466)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.

AUTHORS

Homo sapiens chromosome 18, clone RP11-789C7

TITLE

Unpublished

JOURNAL

2 (bases 1 to 182466)

REFERENCE

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Bouckhalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Headford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

Direct Submission

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 182466)

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Bouckhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Laroque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,


```

unsure      complement(31426..31658)
unsure      /notes="single clone coverage"
unsure      complement(31557..31561)

Query Match      91.6%; Score 17.4; DB 5; Length 182466;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 43399 CTTTACTTCATAGTCTTTG 43381

RESULT 63
LOCUS AC097585/c 182709 bp DNA linear HTG 01-MAR-2002
DEFINITION Sus scrofa clone RP44-35K6, WORKING DRAFT SEQUENCE, 5 ordered
           pieces.
ACCESSION AC097585
VERSION AC097585.2 GI:19033547
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
          Sus.
REFERENCE 1 (bases 1 to 182709)
          Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
          Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
          Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Hegnighi,P.,
          Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q.,
          Legaapi,R., Maduro,Q.L., Maduro,V.B., Masiello,C., Maskeri,B.,
          Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A.,
          Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W.,
          Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
          Young,A., Zhang,L.-H. and Green,E.D.
          NISC Comparative Sequencing Initiative
          Unpublished
REFERENCE 2 (bases 1 to 182709)
          Green,E.D.
          Direct Submission
          Submitted (19-OCT-2001) NIH Intramural Sequencing Center, 8717
          Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 182709)
          Green,E.D.
          Direct Submission
          Submitted (01-MAR-2002) NIH Intramural Sequencing Center, 8717
          Grovemont Circle, Gaithersburg, MD 20877, USA
          On Mar 1, 2002 this sequence version replaced gi:16271923.
          ----- Genome Center
          Center: NIH Intramural Sequencing Center
          Center code: NISC
          Web site: http://www.nisc.nih.gov
          Contact: nisc_mouse@nhgri.nih.gov
          ----- Project Information
          Center project name: ctb
          Center clone name: 035K06

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

```

Assembly program: Phrap; version 0.990319
Consensus quality: 181870 bases at least Q40
Consensus quality: 182166 bases at least Q30
Consensus quality: 182301 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 182309; sum-of-contigs
Quality coverage: 11.37x in Q20 bases; agarose-fp
Quality coverage: 10.29x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1
* 24329: contig of 24329 bp in length
* 24330: gap of unknown length
* 24430: contig of 7193 bp in length
* 31623: gap of unknown length
* 31723: contig of 38989 bp in length
* 70712: gap of unknown length
* 70812: contig of 99994 bp in length
* 70812: 170805: gap of unknown length
* 170806: 182709: contig of 11804 bp in length.
* 170906: 182709: contig of 11804 bp in length.
FEATURES
         Location/Qualifiers
         source
           1..182709
             /organism="Sus scrofa"
             /mol_type="genomic DNA"
             /db_xref="taxon:9823"
             /clone="RP44-35K6"
             /clone_lib="RP44"
             1..38969
               /notes="clone overlaps with GenBank Accession Number
               AC097631 clone RP44-62F17 (center project name cta)"
               1..24329
                 /notes="assembly_fragment"
                 clone_end:T7
                 vector_side:left"
                 24330..24429
                   /estimated_length=unknown
                   24430..31622
                     /notes="assembly_fragment"
                     31623..31722
                       /estimated_length=unknown
                       31723..70711
                         /notes="assembly_fragment"
                         70712..70811
                           /estimated_length=unknown
                           70812..170805
                             /notes="assembly_fragment"
                             108812..182709
                               /notes="clone overlaps with GenBank Accession Number
                               AC097629 clone RP44-265A21 (center project name ctc)"
                               170806..170905
                                 /estimated_length=unknown
                                 170906..182709
                                   /notes="assembly_fragment"
                                   clone_end:SP6
                                   vector_side:right"
         ORIGIN
           Query Match      91.6%; Score 17.4; DB 12; Length 182709;
           Best Local Similarity 94.7%; Pred. No. 3.5e+02;
           Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

           Qy 1 CTTTACTTCATAGTCTTTG 19
               |||||
           Db 135412 CTTTACTTCATAGTCTTTG 135394

```

RESULT 64

AC146837

LOCUS

DEFINITION

AC146837 183404 bp DNA linear HTG 30-OCT-2003
 Canis familiaris clone RP81-48L12, WORKING DRAFT SEQUENCE, 11
 ordered pieces.

ACCESSION

AC146837

VERSION

HTG; HTGS PHASE2; HTGS DRAFT.

KEYWORDS

Canis familiaris (dog)

SOURCE

Canis familiaris

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE

AUTHORS

1 (bases 1 to 183404)
 Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
 Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
 Coleman,H., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R.,
 Karlins,E., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q.,
 Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
 Maskeri,B., McDowell,J., Mullikin,J.C., Paquinigan,C., Pearson,R.,
 Portnoy,M.E., Prasad,A., Reddix-Dugue,N., Schandler,K.,
 Schueler,M.G., Shah,K., Sison,C., Stantripop,S., Thomas,J.W.,
 Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
 Green,E.D.

TITLE

NISC Comparative Sequencing Initiative

REFERENCE

AUTHORS

Green,E.D.

JOURNAL

Direct Submission

TITLE

Submitted (10-OCT-2003) NIH Intramural Sequencing Center, 8717

JOURNAL

Government Circle, Gaithersburg, MD 20877, USA

REFERENCE

AUTHORS

Green,E.D.

JOURNAL

Submitted (30-OCT-2003) NIH Intramural Sequencing Center, 8717

COMMENT

On Oct 30, 2003 this sequence version replaced gi:37620241.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: fyk
 Center clone name: 048L12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 180485 bases at least Q40

Consensus quality: 181505 bases at least Q30

Consensus quality: 182100 bases at least Q20

Insert size: 176000; agarose-fp

Insert size: 182404; sum-of-contigs

Quality coverage: 13.58x in Q20 bases; agarose-fp

Quality coverage: 13.10x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

* This sequence will be replaced

* by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 7223: contig of 7223 bp in length

* 7224: gap of unknown length

* 7324: contig of 23004 bp in length

* 30327: gap of unknown length

* 30328: gap of unknown length

* 30427: contig of 5424 bp in length

* 35851: gap of unknown length

* 35852: contig of 13547 bp in length

* 49498: gap of unknown length

* 49499: contig of 36431 bp in length

* 49529: gap of unknown length

* 86030: contig of 7276 bp in length

* 86130: gap of unknown length

* 93405: contig of 2180 bp in length

* 93406: gap of unknown length

* 136003: contig of 42498 bp in length

* 136103: gap of unknown length

* 136104: contig of 20457 bp in length

* 156560: gap of unknown length

* 156661: contig of 2180 bp in length

* 158841: gap of unknown length

* 158940: contig of 22812 bp in length

* 158941: gap of unknown length

* 181752: contig of 22812 bp in length

* 181753: gap of unknown length

* 181852: contig of 1552 bp in length.

* 181853: contig of 1552 bp in length.

FEATURES

source

1. 183404
 /organism="Canis familiaris"

/mol_type="genomic DNA"

/db_xref="taxon:9615"

/clone_lib="RP81-48L12"

/clone="RP81-48L12"

1. 21343

/note="clone overlaps with GenBank Accession Number
 AC146835 clone RP81-292E2 (center project name fy0)"

misc_feature

1. 7223
 /note="assembly_fragment"

clone_end:SP6

vector_side:left"

7224. 7323

/estimated_length=unknown

7324. 30327

/note="assembly_fragment"

30328. 30427

/estimated_length=unknown

30428. 35851

/note="assembly_fragment"

35852. 35951

/estimated_length=unknown

35952. 49498

/note="assembly_fragment"

49499. 49598

/estimated_length=unknown

49599. 86029

/note="assembly_fragment"

86030. 86129

/estimated_length=unknown

86130. 93405

/note="assembly_fragment"

93406. 93505

/estimated_length=unknown

93506. 136003

/note="assembly_fragment"

136004. 136103

/estimated_length=unknown

136104. 156560

/note="assembly_fragment"

156561. 156660

/estimated_length=unknown

156661. 158840

```

/notes="assembly_fragment"
158841..158940
/estimated_length=unknown
158941..181752
/notes="assembly_fragment"
166557..183404
/notes="clone overlaps with GenBank Accession Number
AC146771 clone RP81-121B12 (center project name fyl)"
181753..181852
/estimated_length=unknown
181853..183404
/notes="assembly_fragment
clone_end:T7
vector_side:right"

ORIGIN
Query Match          91.6%; Score 17.4; DB 12; Length 183404;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTGG 19
|||||
Db 165279 CATTACTTCATAGCTCTTGG 165297

RESULT 65
AC011947/c
LOCUS
DEFINITION Homo sapiens clone RP11-17D12, WORKING DRAFT SEQUENCE, 6 unordered
pieces.
ACCESSION AC011947
VERSION AC011947.4 GI:14190733
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 183812)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-17D12
Unpublished
2 (bases 1 to 183812)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castelle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 23, 2001 this sequence version replaced gi:7387330.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3625
Center clone name: 17_D12
----- Summary Statistics

```

```

Sequencing vector: M13; M7815; 0% of reads
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 179876 bases at least Q40
Consensus quality: 182173 bases at least Q30
Consensus quality: 182976 bases at least Q20
Insert size: 183000; agarose-fp
Insert size: 183312; sum-of-contigs
Quality coverage: 7.7 in Q20 bases; agarose-fp
Quality coverage: 7.7 in Q20 ba.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 81681: contig of 81681 bp in length
* 81682 81781: gap of 100 bp
* 81782 83362: contig of 1581 bp in length
* 83363 83462: gap of 100 bp
* 83463 84806: contig of 1344 bp in length
* 84807 85699: gap of 100 bp
* 84907 85799: gap of 100 bp
* 85800 115070: contig of 29271 bp in length
* 115071 115170: gap of 100 bp
* 115171 183812: contig of 68642 bp in length.
FEATURES
source
1..183812
/organism="Homo sapiens"
/db type="genomic DNA"
/db xref="taxon:9606"
/clone="RP11-17D12"
/clone lib="RPC1-11 Human Male BAC"
misc_feature
1..81681
/notes="assembly_fragment
clone_end:SP6
vector_side:left"
gap
81682..81781
/estimated_length=100
misc_feature
81782..83362
/notes="assembly_fragment"
gap
83363..83462
/estimated_length=100
misc_feature
83463..84806
/notes="assembly_fragment"
gap
84807..84906
/estimated_length=100
misc_feature
84907..85699
/notes="assembly_fragment"
gap
85700..85799
/estimated_length=100
misc_feature
85800..115070
/notes="assembly_fragment"
gap
115071..115170
/estimated_length=100
misc_feature
115171..183812
/notes="assembly_fragment
clone_end:T7
vector_side:right"
ORIGIN

```

```

Query Match          91.6%; Score 17.4; DB 12; Length 183812;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTGG 19
|||||
Db 100430 CTTTACTTCATAGCTCTTGG 100412

```

```

RESULT 66
AC069476      184824 bp      DNA      linear      PRI 09-APR-2003
LOCUS      Homo sapiens chromosome 11 clone b430h11 map 11, complete sequence.
DEFINITION
AC069476
VERSION      AC069476.25 GI:29244703
KEYWORDS
SOURCE      HTG.
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 184824)
Deschamps,S., Richter,N. and Roe,B.A.
Homo sapiens Chromosome 11 BAC Clone b430h11
UNPUBLISHED
2 (bases 1 to 184824)
Deschamps,S., Richter,N. and Roe,B.A.
Direct Submission
Submitted (01-JUN-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 184824)
Deschamps,S., Richter,N. and Roe,B.A.
Direct Submission
Submitted (14-MAR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 184824)
Deschamps,S., Richter,N. and Roe,B.A.
Direct Submission
Submitted (19-MAR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 184824)
Deschamps,S., Richter,N. and Roe,B.A.
Direct Submission
Submitted (26-MAR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 184824)
Deschamps,S., Richter,N. and Roe,B.A.
Direct Submission
Submitted (09-APR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Mar 26, 2003 this sequence version replaced gi:28951189.
-----
Center: Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source
1. 184824
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clones="b430h11"
/clone_libs="RPCI Human Female PAC Library 6"
/note="This clone was originally contaminated with 10kb of
mollusc sequence"
ORIGIN
Query Match      91.6%; Score 17.4; DB 5; Length 184824;
Beat Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGCTTTG 19
Db 50197 CTTTACTTCATGCTTTG 50215

```

RESULT 67

AC132501

LOCUS

DEFINITION

AC132501

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AC132501 185079 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH230-306G10, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.

AC132501
AC132501.3 GI:25139146
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 185079)

Muzny,D., Maric, M., Metzker, M., Lee, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, J., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhera, L., Louissegh, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaekemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanai, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 185079)

Rat Genome Sequencing Consortium.
Direct Submission

Submitted (01-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

```

REFERENCE
AUTHORS 3 (bases 1 to 185079)
TITLE Rat Genome Sequencing Consortium.
JOURNAL Direct Submission
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 20, 2002 this sequence version replaced gi:22855833.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBSN
Center clone name: CH230-306G10
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 173513 bases at least Q40
Consensus quality: 175951 bases at least Q30
Consensus quality: 177426 bases at least Q20
Estimated insert size: 176985; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 29987: contig of 29987 bp in length
* 29988 30087: gap of unknown length
* 30088 183113: contig of 153026 bp in length
* 183114 183213: gap of unknown length
* 183214 185079: contig of 1866 bp in length.
*
* Location/Qualifiers
* 1..185079
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-306G10"
* 29988..30087
* /estimated_length=unknown
* 183114..183213
* /estimated_length=unknown
*
* gap
*
* gap
*
* ORIGIN
Query Match 91.6%; Score 17.4; DB 12; Length 185079;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| ||||| |||||
Db 147719 CTTTACTTCATAGTCTTTG 147737

FEATURES
source
1..185079
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

gap
103107: contig of 103107 bp in length
103108 103207: gap of unknown length
103208 114179: contig of 10972 bp in length
114180 114279: gap of unknown length
114280 117610: contig of 3331 bp in length
117611 117710: gap of unknown length
117711 120716: contig of 3006 bp in length
120717 120816: gap of unknown length
120817 186281: contig of 65465 bp in length.
1..186281
/organism="Mus musculus"
/mol_type="genomic DNA"

LOCUS 186281 bp DNA linear HTG 12-MAY-2005
Mus musculus chromosome 1 clone RP23-456J17 map 1, *** SEQUENCING
IN PROGRESS ***, 5 unordered pieces.
AC161349 1 GI:63253447
HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
Mus musculus (house mouse)
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 186281)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP23-456J17
Unpublished
2 (bases 1 to 186281)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorr, L.,
Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karada, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Liu, A., Mabbitt, R.,
MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (12-MAY-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seg.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L31746
Center clone name: 456_J_17
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 103107: contig of 103107 bp in length
* 103108 103207: gap of unknown length
* 103208 114179: contig of 10972 bp in length
* 114180 114279: gap of unknown length
* 114280 117610: contig of 3331 bp in length
* 117611 117710: gap of unknown length
* 117711 120716: contig of 3006 bp in length
* 120717 120816: gap of unknown length
* 120817 186281: contig of 65465 bp in length.
1..186281
/organism="Mus musculus"
/mol_type="genomic DNA"

FEATURES
source
1..186281
/organism="Mus musculus"
/mol_type="genomic DNA"

```



```

/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-456J17"
/clone_lib="RPCI-23 Female Mouse BAC"
103108..103207
/estimated_length=unknown
114180..114279
/estimated_length=unknown
117611..117710
/estimated_length=unknown
120717..120816
/estimated_length=unknown

ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 186281;
Best Local Similarity 94.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 186103 CTTTACTTCATGGTCTTTG 186085

RESULT 70
AC151368/c
LOCUS AC151368 189563 bp DNA linear PRI 04-NOV-2005
DEFINITION Aotus nancymaae clone CH258-251019, complete sequence.
ACCESSION AC151368
VERSION AC151368.3 GI:79758574
KEYWORDS HTG.
SOURCE Aotus nancymaae (Ma's night monkey)
ORGANISM Aotus nancymaae
Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
Cebidae; Aotinae; Aotus.
1 (bases 1 to 189563)
Antellis,A., Ayele,K., Bass,D., Benjamin,B., Bera,J., Chu,G.,
Blakesley,R.W., Bouffard,G., Brinkley,C., Brooks,S., Guan,X.,
Coleman,H., Franks,S., Fukuoka,T., Gestole,M., Greene,A., Han,J.,
Gupta,J., Gurson,N., Haghighi,P., Han,E., Han,J., Hansen,N.,
Ho,S.-L., Hu,P., Hunter,G., Hurler,B., Idol,J.R., Kwong,P.,
Laric,P., Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M.,
Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B.,
McDowell,J., Montemayor,C., Mullikin,J.C., Park,M., Portnoy,M.E.,
Prasad,A., Puri,O., Rantzi,K., Reddix-Dugue,N., Sante,A.,
Schandler,K., Schueler,M.G., Sison,C., Stantropop,S., Tave,A.,
Thomas,J.W., Thomas,P.J., Tsipouri,V., Ung,L., Vogt,J.L.,
Wetherby,K.D., Withers,T.R., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 189563)
Green,E.D.
Direct Submission
Submitted (25-AUG-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 189563)
Green,E.D.
Direct Submission
Submitted (14-SEP-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
4 (bases 1 to 189563)
Green,E.D.
Direct Submission
Submitted (04-NOV-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Nov 4, 2005 this sequence version replaced gi:52000558.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: hzf
Center clone name: 251019

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

```

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features section.

FEATURES

source

Location/Qualifiers

1. .189563
 /organism="Aotus nancymae"
 /mol_type="genomic DNA"
 /db_xref="taxon:37293"
 /clone="CH258-251O19"
 /clone_lib="CH258"
 /note="BAC resource: <http://bacpac.chori.org/>"

ORIGIN

Query Match 91.6%; Score 17.4; DB 5; Length 189563;
 Best Local Similarity 94.7%; Pred. No. 3.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
 |||||

Db 187640 CTTTACTTCATAGCTCTG 187622

RESULT 71

AP001797

LOCUS

DEFINITION Homo sapiens chromosome 18 clone RP11-748120 map 18q12, WORKING
 DRAFT SEQUENCE, 27 unordered pieces.

ACCESSION

AP001797

VERSION

AP001797.2 GI:8117469

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

1 (bases 1 to 190550)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 190,550 genomic DNA of 18q12

Published Only in DataBase (2000)

2 (bases 1 to 190550)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (11-APR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gsr.riken.go.jp,
 URL:<http://hgp.gsc.riken.go.jp/>, Tel:81-42-778-9923,
 Fax:81-42-778-9924)

On May 30, 2000 this sequence version replaced gi:7592912.

COMMENT

----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: <http://hgp.gsc.riken.go.jp/>

Contact: hattori@gsr.riken.go.jp

----- Project Information

Center project name: HumDraft18

Center clone name: RP11-748120

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990129

Consensus quality: 178641 bases at least Q40

Consensus quality: 184743 bases at least Q30

Insert size: 187950; sum-of-contigs

Quality coverage: 4.84x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 27 contigs. The true order of the pieces is not known and their

order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

1 16718 contig of 16718 bp in length
 16819 30449 contig of 13631 bp in length
 30550 45067 contig of 15058 bp in length
 45708 59752 contig of 14045 bp in length
 59853 75894 contig of 16042 bp in length
 75995 87034 contig of 11040 bp in length
 87135 97062 contig of 9928 bp in length
 97163 106435 contig of 9273 bp in length
 106536 114800 contig of 8265 bp in length
 114901 121600 contig of 6700 bp in length
 121701 129422 contig of 7722 bp in length
 129523 138873 contig of 9350 bp in length
 138973 145365 contig of 6393 bp in length
 145466 151207 contig of 5742 bp in length
 151308 155673 contig of 4366 bp in length
 155774 160475 contig of 5105 bp in length
 160576 169630 contig of 3850 bp in length
 169731 174769 contig of 5039 bp in length
 174870 178072 contig of 3203 bp in length
 178173 180351 contig of 2179 bp in length
 180452 181295 contig of 844 bp in length
 181396 184552 contig of 3157 bp in length
 184653 186282 contig of 1630 bp in length
 186383 188196 contig of 1814 bp in length
 188397 189386 contig of 1090 bp in length
 189487 190550 contig of 1064 bp in length

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 27 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 16718: contig of 16718 bp in length
 16819 16818: gap of 100 bp
 16819 30449: contig of 13631 bp in length
 30450 30549: gap of 100 bp
 30550 45607: contig of 15058 bp in length
 45708 45707: gap of 100 bp
 45708 59752: contig of 14045 bp in length
 59753 59852: gap of 100 bp
 59853 75894: contig of 16042 bp in length
 75895 75994: gap of 100 bp
 75995 87134: contig of 11040 bp in length
 87135 87133: gap of 100 bp
 87135 97062: contig of 9928 bp in length
 97063 97162: gap of 100 bp
 97163 106435: contig of 9273 bp in length
 106436 106536: gap of 100 bp
 106536 114800: contig of 8265 bp in length
 114801 114900: gap of 100 bp
 114901 121600: contig of 6700 bp in length
 121601 121700: gap of 100 bp
 121701 129422: contig of 7722 bp in length
 129423 129522: gap of 100 bp
 129523 138872: contig of 9350 bp in length
 138873 138972: gap of 100 bp
 138973 145365: contig of 6393 bp in length
 145366 145465: gap of 100 bp
 145466 151207: contig of 5742 bp in length
 151208 151307: gap of 100 bp
 151308 155673: contig of 4366 bp in length
 155674 155773: gap of 100 bp
 155774 160475: contig of 4702 bp in length
 160476 160575: gap of 100 bp

* 160576 165680: contig of 5105 bp in length
 * 165780: gap of 100 bp
 * 165781 169630: contig of 3850 bp in length
 * 169730: gap of 100 bp
 * 169731 174769: contig of 5039 bp in length
 * 174770 174869: gap of 100 bp
 * 174870 178072: contig of 3203 bp in length
 * 178073 178172: gap of 100 bp
 * 178173 180351: contig of 2179 bp in length
 * 180352 180451: gap of 100 bp
 * 180452 181295: contig of 844 bp in length
 * 181296 181396: gap of 100 bp
 * 181397 184552: contig of 3157 bp in length
 * 184553 184652: gap of 100 bp
 * 184653 186282: contig of 1630 bp in length
 * 186283 186382: gap of 100 bp
 * 186383 188196: contig of 1814 bp in length
 * 188197 188296: gap of 100 bp
 * 188297 189386: contig of 1090 bp in length
 * 189387 189486: gap of 100 bp
 * 189487 190550: contig of 1064 bp in length.

FEATURES

source

1. 190550
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"

/map="18q12"

/clone="RP11-748120"

1. 16718

/note="assembly_fragment"

16819. 30449

/note="assembly_fragment"

30550. 45607

/note="assembly_fragment"

45708. 59752

/note="assembly_fragment"

59853. 75894

/note="assembly_fragment"

75995. 87034

/note="assembly_fragment"

87135. 97062

/note="assembly_fragment"

97163. 106435

/note="assembly_fragment"

106536. 114800

/note="assembly_fragment"

114901. 121600

/note="assembly_fragment"

121701. 129422

/note="assembly_fragment"

129523. 138872

/note="assembly_fragment"

138973. 145365

/note="assembly_fragment"

145466. 151207

/note="assembly_fragment"

151308. 155673

/note="assembly_fragment"

155774. 160475

/note="assembly_fragment"

160576. 165680

/note="assembly_fragment"

165781. 169630

/note="assembly_fragment"

169731. 174769

/note="assembly_fragment"

174870. 178072

/note="assembly_fragment"

178173. 180351

/note="assembly_fragment"

180452. 181295

/note="assembly_fragment"

misc_feature

181396. 184552

/note="assembly_fragment"

misc_feature

184653. 186282

/note="assembly_fragment"

misc_feature

186383. 188196

/note="assembly_fragment"

misc_feature

188297. 189386

/note="assembly_fragment"

Query Match 91.6%; Score 17.4; DB 12; Length 190550;

Best Local Similarity 94.7%; Pred. No. 3.6e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGTCTTTG 19

|||||

Db 70450 CTTTACTTCATAGTCTTGG 70468

RESULT 72

AC097625/c

LOCUS

AC097625 194492 bp DNA linear PRI 29-MAY-2002
 Homo sapiens X BAC RP11-478H11 (Roswell Park Cancer Institute Human
 BAC Library) complete sequence.

ACCESSION AC097625 AC058807

VERSION AC097625.11 GI:21240506

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 194492)

REFERENCE

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Albrooks, S.L., Amarutunge, H.C., Are, J.R., Banks, T., Barbaria, J.,
 Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J.,
 Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F.,
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
 Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
 Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
 Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
 Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
 Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
 Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
 Loulseg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
 Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
 Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
 Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.
 Direct Submission
 TITLE
 JOURNAL
 REFERENCE
 2 (bases 1 to 194492)

```
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 194492)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 194492)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT
On May 29, 2002 this sequence version replaced gi:21206073.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
of STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES
source
1..194492
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/Chromosome="X"
/Clone="RP11-478H11"
88..242
repeat_region /rpt_family="AluJo/FRAM"
648..2719
repeat_region /rpt_family="L1MA4"
2720..3012
repeat_region /rpt_family="AluSx"
3013..4086
repeat_region /rpt_family="L1MA4"
4100..4396
repeat_region /rpt_family="AluJo"

complement(4786..5095)
/rpt_family="AluSg"
6104..6764
/rpt_family="L1ME3A"
6869..6980
/rpt_family="L2"
7028..7075
/rpt_family="(CA)n"
7076..7106
/rpt_family="(CA)n"
7178..7260
/rpt_family="(TA)n"
8004..8054
/rpt_family="(TTTC)n"
complement(8061..8194)
/rpt_family="FLAM_C"
8961..9258
/rpt_family="AluY"
11398..11450
/rpt_family="AT rich"
complement(11453..11520)
/rpt_family="MIR"
complement(11596..12790)
/rpt_family="L2"
12833..13015
/rpt_family="GA-rich"
13040..13661
/rpt_family="GA-rich"
13699..13881
/rpt_family="GA-rich"
14315..14612
/rpt_family="AluSx"
14616..14661
/rpt_family="(GAAA)n"
16525..16625
/rpt_family="MIR"
16850..17414
/rpt_family="LTR55"
17828..17950
/rpt_family="MIR"
18539..18575
/rpt_family="MIR"
18576..18881
/rpt_family="MER33"
18882..18922
/rpt_family="MIR"
complement(19026..19064)
/rpt_family="L2"
complement(19064..19207)
/rpt_family="MIR"
19371..19397
/rpt_family="(TC)n"
complement(19726..20140)
/rpt_family="L1ME"
21404..22846
/rpt_family="Tigger1"
22874..22947
/rpt_family="(TTAA)n"
22968..23906
/rpt_family="Tigger1"
24713..24922
/rpt_family="L1PA6"
complement(25204..25514)
/rpt_family="AluY"
25758..25949
/rpt_family="MER30"
26165..26240
/rpt_family="L2"
complement(26692..27003)
/rpt_family="AluY"
28270..28398
/rpt_family="L2"
complement(29047..29258)
```

```

Query Match          91.6%; Score 17.4; DB 5; Length 194492;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| |||||
Db 29223 CTTTACTCCATAGTCTTTG 29205

RESULT 73
LOCUS AC122492/c
DEFINITION AC122492 196275 bp DNA linear ROD 13-NOV-2003
          Mus musculus BAC clone RP24-390N7 from chromosome 5, complete
          sequence.
VERSION AC122492.3 GI:28209796
KEYWORDS
SOURCE HTG.
ORGANISM Mus musculus (house mouse)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 196275)
          Harkins, R., Cotton, M., Haglund, K. and Bielicki, L.
          The sequence of Mus musculus BAC clone RP24-390N7
          Unpublished (2001)
REFERENCE 2 (bases 1 to 196275)
          Wilson, R.
          Sequencing of Mus musculus
          Unpublished (2001)
REFERENCE 3 (bases 1 to 196275)
          McPherson, J.D. and Waterston, R.H.
          Direct Submission
          Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 196275)
          McPherson, J.D. and Waterston, R.H.
          Direct Submission
          Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 196275)
          McPherson, J.D. and Waterston, R.H.
          Direct Submission
          Submitted (05-FEB-2003) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 196275)
          Wilson, R.
          Direct Submission
          Submitted (13-NOV-2003) Department of Genetics, Washington
          University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
          On Feb 5, 2003 this sequence version replaced gi:21539185.
          ----- Genome Center
          Center: Washington University Genome Sequencing Center
          Center code: WUGSC
          Web site: http://genome.wustl.edu
          Contact: submissions@watson.wustl.edu
          ----- Summary Statistics
          -----
          Center project name: M_BB0390N07
          -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC115118.

FEATURES	Location/Qualifiers	source
repeat_region	38..369	/rpt_family="MaLR"
repeat_region	370..964	/rpt_family="L1"
repeat_region	963..1419	/rpt_family="L1"
repeat_region	2316..2473	/rpt_family="L1"
repeat_region	4038..4089	/rpt_family="B4"
repeat_region	4104..4214	/rpt_family="RMR1B"
repeat_region	4394..4523	/rpt_family="Alu"
repeat_region	4798..5009	/rpt_family="B2"
repeat_region	5154..5333	/rpt_family="B2"
repeat_region	7461..8097	/rpt_family="L1"
repeat_region	8534..9541	/rpt_family="L1"
repeat_region	9578..9837	/rpt_family="L1"
repeat_region	9860..10091	/rpt_family="L1"
repeat_region	10092..10518	/rpt_family="ERVK"
repeat_region	10519..10703	/rpt_family="L1"
repeat_region	11526..12174	/rpt_family="L1"
repeat_region	12559..13158	/rpt_family="L1"
repeat_region	13158..13418	/rpt_family="L1"
repeat_region	13407..13834	/rpt_family="L1"
repeat_region	13831..14876	/rpt_family="L1"
repeat_region	14877..15272	/rpt_family="MaLR"
repeat_region	15273..15311	/rpt_family="L1"
repeat_region	15310..16540	/rpt_family="L1"
repeat_region	16944..17274	/rpt_family="L1"

```
repeat_region 17275..17459
/rpt_family="B2"
tRNA 17280..17353
/product="tRNA-Ser"
/notes="Likely pseudogene (HMM SC=42.36 / Sec struct
SC=22.25)"
repeat_region 17460..17685
/rpt_family="L1"
repeat_region 17686..17784
/rpt_family="L1"
repeat_region 17799..20144
/rpt_family="MaLR"
repeat_region 20913..21432
/rpt_family="ERVK"
repeat_region 22460..22804
/rpt_family="MaLR"
repeat_region 24845..25333
/rpt_family="L1"
repeat_region 25777..25827
/rpt_family="ERV1"
repeat_region 26215..26709
/rpt_family="L1"
repeat_region 26745..27207
/rpt_family="L1"
repeat_region 29209..29949
/rpt_family="L1"
repeat_region 31048..31222
/rpt_family="B2"
repeat_region 32006..32407
/rpt_family="L1"
repeat_region 32442..32788
/rpt_family="L1"
repeat_region 33915..34069
/rpt_family="B4"
repeat_region 34573..35037
/rpt_family="L1"
repeat_region 35872..35927
/rpt_family="ERV1"
repeat_region 36012..36163
/rpt_family="L1"
repeat_region 36323..36553
/rpt_family="L1"
repeat_region 37016..37126
/rpt_family="Alu"
repeat_region 37425..37616
/rpt_family="MER1_type"
repeat_region 39478..39609
/rpt_family="ID"
repeat_region 40439..40836
/rpt_family="L1"
repeat_region 40994..41098
/rpt_family="L1"
repeat_region 41197..41448
/rpt_family="MaLR"
repeat_region 43283..43551
/rpt_family="L1"
repeat_region 43790..43979
/rpt_family="B4"
repeat_region 44009..44154
/rpt_family="Alu"
repeat_region 44156..44271
/rpt_family="B4"
repeat_region 44664..45063
/rpt_family="ERVK"
repeat_region 48337..49516
/rpt_family="ERVK"
Query Match 91.6%; Score 17.4; DB 6; Length 196275;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 5924 CTTTCTTCATAGTCTTTG 5906
196700 bp DNA linear HTG 13-MAR-2002
Sus scrofa clone RP44-265A21, WORKING DRAFT SEQUENCE, 5 ordered
pieces.
AC097629
AC097629/c
LOCUS AC097629.2 GI:19387618
DEFINITION HTG: HTGS PHASE2; HTGS_DRAFT.
VERSION HTG: HTGS PHASE2; HTGS_DRAFT.
KEYWORDS Sus scrofa (pig)
SOURCE Sus scrofa
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE 1 (bases 1 to 196700)
AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghghi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,
Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Masiello, C., Maskeri, B., Mastrian, S.D.,
McCloskey, J.C., McDowell, J., Paguirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantropop, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
TITLE 2 (bases 1 to 196700)
JOURNAL Green, E.D.
REFERENCE Direct Submission
AUTHORS Submitted (20-OCT-2001) NIH Intramural Sequencing Center, 8717
TITLE Grovmont Circle, Gaithersburg, MD 20877, USA
JOURNAL Grovmont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 196700)
AUTHORS Green, E.D.
JOURNAL Direct Submission
TITLE Submitted (13-MAR-2002) NIH Intramural Sequencing Center, 8717
JOURNAL Grovmont Circle, Gaithersburg, MD 20877, USA
COMMENT On Mar 13, 2002 this sequence version replaced gi:16303428.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@hgrl.nih.gov
----- Project Information
Center project name: ctc
Center clone name: 265A21
```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 195502 bases at least Q40
Consensus quality: 196012 bases at least Q30
Consensus quality: 196230 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 196300; sum-of-contigs
Quality coverage: 10.63x in Q20 bases; agarose-fp
Quality coverage: 9.42x in Q20 bases; sum-of-contigs

Qy 1 CTTTCTTCATAGTCTTTG 19
|||||

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 61967: contig of 61967 bp in length

* 61968 62067: gap of unknown length

* 62068 128700: contig of 66633 bp in length

* 128701 128800: gap of unknown length

* 128801 144849: contig of 16049 bp in length

* 144850 144949: gap of unknown length

* 144950 169039: contig of 24090 bp in length

* 169040 169139: gap of unknown length

* 169140 196700: contig of 27561 bp in length.

FEATURES

source

1. 196700
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="RP44-265A21"
/clone="RP44"

misc_feature

1. 73532
/note="clone overlaps with GenBank Accession Number AC097585 clone RP44-35K6 (center project name ctb)"

misc_feature

1. 61967
/note="assembly_fragment"
clone end:T7
vector_side:left"

gap

61968..62067

misc_feature

/estimated_length=unknown

misc_feature

62068..128700
/note="assembly_fragment"

misc_feature

123997..196700
/note="clone overlaps with GenBank Accession Number AC098815 clone RP44-213D24 (center project name ctd)"

gap

128701..128800

misc_feature

/estimated_length=unknown

gap

128801..144849

misc_feature

144850..144949
/estimated_length=unknown

gap

144950..169039
/note="assembly_fragment"

misc_feature

169040..169139
/estimated_length=unknown

misc_feature

169140..196700
/note="assembly_fragment"

misc_feature

clone end:SP6
vector_side:right"

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 196700;

Best Local Similarity 94.7%; Pred. No. 3.6e+02; Length 196700;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CTTTACTTCATAGCTTTG 19

26597 CTTTACTTCATCTTTG 26579

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 197439)

Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 1, clone RP24-442E5

Unpublished

2 (bases 1 to 197439)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,

Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,

Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,

Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,

Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,

Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,

Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Ramond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 197439)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,

Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,

Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,

Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,

DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,

Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,

Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,

Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,

Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,

MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,

Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,

Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,

Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,

Zimmer,A. and Zody,M.

Direct Submission

Submitted (18-MAR-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 197439)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,

Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,

Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,

Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,

DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,

Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,

Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,

Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,

Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,

MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,

Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,

Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,

Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,

Zimmer,A. and Zody,M.

Direct Submission

Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

TITLE

Direct Submission

Submitted (29-APR-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Apr 29, 2005 this sequence version replaced gi:61651929.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L19314

Center clone name: 442_E_5

FEATURES

source

1. .197439

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="1"

/map="1"

/clone="RP24-442E5"

/clone_lib="RPC1-24 Male Mouse BAC"

complement(1. .1725)

/rpt_family="L1"

1. 634

/notes="PCR product sequence only"

1. .112

/notes="single clone coverage"

complement(582. .634)

/notes="single clone coverage"

complement(1422. .1527)

/notes="single clone coverage"

complement(1724. .1782)

/rpt_family="L1VL2"

/rpt_family="L1 MM"

complement(1778. .8773)

/rpt_family="L1 MM"

complement(8775. .9588)

/rpt_family="L1VL2"

complement(9694. .9762)

/rpt_family="RMER6B"

10532. .11597

/rpt_family="MULV-int"

complement(10593. .10660)

/notes="single clone coverage"

complement(10800. .10864)

/notes="single clone coverage"

complement(10817. .10826)

/notes="<30 Qual SNGL region"

11743. .11933

/rpt_family="MULV-int"

12488. .13317

/rpt_family="MULV-int"

13594. .13997

/rpt_family="MULV-int"

14659. .14685

/rpt_family="(TG)n"

15285. .15370

/rpt_family="(TRCC)n"

complement(15310. .15545)

/notes="single clone coverage"

complement(15371. .15557)

/rpt_family="B2_Mm2"

16901. .16930

repeat_region
 /rpt_family="AT_rich"
 18474. .18533
 /rpt_family="(CATA)n"
 complement(18536. .18626)
 repeat_region
 /rpt_family="L1MC4"
 19529. .19554
 /rpt_family="(TG)n"
 20605. .20648
 /rpt_family="(TTTTG)n"
 complement(20661. .20802)
 repeat_region
 /rpt_family="B1_MM"
 21732. .21765
 /rpt_family="(CA)n"
 21770. .21836
 /rpt_family="(TATATG)n"
 22515. .22654
 /rpt_family="B1P"
 22657. .25314
 /rpt_family="L1_MM"
 complement(26335. .26508)
 repeat_region
 /rpt_family="B3"
 complement(26723. .32928)
 repeat_region
 /rpt_family="L1_MM"
 27213. .27245
 /notes="single clone coverage"
 27216. .27221
 /notes="<30 Qual SNGL region"
 34104. .34158
 /rpt_family="(CA)n"
 complement(34159. .34472)
 repeat_region
 /rpt_family="MTD"
 34507. .34840
 /rpt_family="IAPLTR1_MM"
 34907. .37169
 /rpt_family="IAPLTR1_MM-int"
 37169. .37224
 /rpt_family="IAPLTR1_MM-int"
 37223. .37579
 /rpt_family="IAPLTR1_MM-int"
 37571. .38345
 /rpt_family="IAPLTR1_MM-int"
 38348. .38681
 /rpt_family="IAPLTR1_MM"
 complement(38752. .39726)
 repeat_region
 /rpt_family="Lx8"
 complement(40083. .40510)
 repeat_region
 /rpt_family="L1MA5"

Query Match 91.6%; Score 17.4; DB 6; Length 197439;

Best Local Similarity 94.7%; Pred. No. 3.6e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 146317 CTTTCTTCATAGCTTTG 146335

RESULT 76

AC115118 198448 bp DNA linear ROD 05-NOV-2003
 LOCUS Mus musculus BAC clone RP23-84N24 from 5, complete sequence.
 DEFINITION AC115118
 ACCESSION AC115118
 VERSION AC115118.3 GI:22138690
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 198448)

AUTHORS Abbott, S., Haakenson, W. and Doebber, A.

TITLE The sequence of Mus musculus BAC clone RP23-84N24

JOURNAL Unpublished (2001)

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 210493;
 Best Local Similarity 94.7%; Pred. No. 3.6e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 24120 CTTTACTTCATAGTCCTG 24102
 |||||

RESULT 78

AC113245/c
 LOCUS
 DEFINITION Mus musculus chromosome 7, clone RP23-175D5, complete sequence.
 AC113245
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE HTG.
 Mus musculus
 Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 213428)
 Birren, B., Nusbaum, C. and Lander, E.

Mus musculus chromosome 7, clone RP23-175D5

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 213428)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Riese, C., Rogov, P., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (27-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 213428)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mieng, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

TITLE

Submitted (20-DEC-2005) Broad Institute of MIT and Harvard, 320
 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 213428)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mieng, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

Center: Broad Institute of MIT and Harvard

Center code: W1BR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L23486

Center clone name: 175_D_5

----- Location/Qualifiers

1. 213428

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE


```

/db_xref="taxon:9483"
/clone="CH259-225J7"
/clone_lib="CH259"
/note="BAC resource: http://bacpac.chori.org/"
1938..31985
/note="PCR product sequence only"
misc_feature
ORIGIN

```

```

Query Match          91.6%; Score 17.4; DB 5; Length 216802;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 143520 CTTTACTTCATAGTCTG 143502

```

```

RESULT 80
AC147845/c
LOCUS
DEFINITION Saimiri boliviensis boliviensis clone CH254-335K11, WORKING DRAFT
ACCESSION AC147845
VERSION AC147845.2 GI:42538994
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Saimiri boliviensis boliviensis (Bolivian squirrel monkey)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
Cebidae; Cebinae; Saimiri.
1 (bases 1 to 217669)
Antonelias,A., Avelle,K., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,J.R., Jones,C., Karlins,B., Kim,H., Kwong,P., Laric,P.,
Larson,S., Lee-lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddi-Dugue,N., Schander,K., Schueler,M.G., Shah,K., Sison,C.,
Scantripp,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 217669)
Green,E.D.
Direct Submission
Submitted (15-JAN-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 217669)
Green,E.D.
Direct Submission
Submitted (12-FEB-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Feb 12, 2004 this sequence version replaced gi:40882645.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: GUD
Center clone name: 335K11

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out

Gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

```

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 215445 bases at least Q40
Consensus quality: 216320 bases at least Q30
Consensus quality: 216749 bases at least Q20
Insert size: 203000; agarose-fp
Insert size: 217069; sum-of-contigs
Quality coverage: 9.44x in Q20 bases; agarose-fp
Quality coverage: 8.83x in Q20 bases; sum-of-contigs
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

```

1 51328: contig of 51328 bp in length
51329 51428: gap of unknown length
51429 134408: contig of 82980 bp in length
134409 134508: gap of unknown length
134509 178461: contig of 43953 bp in length
178462 178562: gap of unknown length
178563 184771: contig of 6210 bp in length
184772 184871: gap of unknown length
184872 190259: contig of 5388 bp in length
190260 190359: gap of unknown length
190360 216319: contig of 25960 bp in length
216320 216419: gap of unknown length
216420 217669: contig of 1250 bp in length.

```

FEATURES

```

source
1..217669
/organism="Saimiri boliviensis boliviensis"
/mol_type="genomic DNA"
/sub_species="boliviensis"
/db_xref="taxon:94832"
/clone="CH254-335K11"
/clone_lib="CH254"
/note="BAC resource: http://bacpac.chori.org/"
1..51328
/note="assembly_fragment"
clone end:T7
vector_side:left"
51329..51428
/estimated_length=unknown
51429..134408
/note="assembly_fragment"
134409..134508
/estimated_length=unknown
134509..178461
/note="assembly_fragment"
178462..178561
/estimated_length=unknown
178562..184771
/note="assembly_fragment"
184772..184871
/estimated_length=unknown
184872..190259
/note="assembly_fragment"
190260..190359
/estimated_length=unknown
190360..216319
/note="assembly_fragment"
200723..217669
/note="clone overlaps with GenBank Accession Number
AC147932 clone CH254-331B16 (center project name guc)"

```

misc_feature

```
1..51328
```

```
/note="assembly_fragment"
```

```
clone end:T7
```

```
vector_side:left"
```

```
51329..51428
```

```
/estimated_length=unknown
```

```
51429..134408
```

```
/note="assembly_fragment"
```

```
134409..134508
```

```
/estimated_length=unknown
```

```
134509..178461
```

```
/note="assembly_fragment"
```

```
178462..178561
```

```
/estimated_length=unknown
```

```
178562..184771
```

```
/note="assembly_fragment"
```

```
184772..184871
```

```
/estimated_length=unknown
```

```
184872..190259
```

```
/note="assembly_fragment"
```

```
190260..190359
```

```
/estimated_length=unknown
```

```
190360..216319
```

```
/note="assembly_fragment"
```

```
200723..217669
```

```
/note="clone overlaps with GenBank Accession Number
```

```
AC147932 clone CH254-331B16 (center project name guc)"
```

```

gap          216320..216419
             /estimated_length=unknown
misc_feature 216420..217669
             /note="assembly_fragment
             clone end:SP6
             vector_side:right"

ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 217669;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCCTTGG 19
        |||||
Db      135502 CTTTACTTCATAGTCCTG 135484

RESULT 81
AC092434/c
LOCUS       AC092434                218872 bp DNA linear HTG 27-JUL-2001
DEFINITION Homo sapiens chromosome 4 clone RP11-210L21, *** SEQUENCING IN
            PROGRESS ***, 40 unordered pieces.
ACCESSION   AC092434
VERSION     AC092434.1 GI:145959556
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE   1 (bases 1 to 218872)
            Waterston,R.H.
            The sequence of Homo sapiens clone
            Unpublished
REFERENCE   2 (bases 1 to 218872)
            Waterston,R.H.
            Direct Submission
            Submitted (04-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0210L21
----- Summary Statistics -----
Sequencing vector: M13; 14%
Sequencing vector: plasmid; 86%
Chemistry: Dye-primer ET; 14% of reads
Chemistry: Dye-terminator Big Dye; 86% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 188006 bases at least Q40
Consensus quality: 196601 bases at least Q30
Consensus quality: 202813 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 214972; sum-of-contigs
Quality coverage: 7.25 in Q20 bases; agarose-fp
Quality coverage: 5.64 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2669: contig of 2669 bp in length
* 2670 2769: gap of unknown length
* 2770 4554: contig of 1785 bp in length
* 4555 4654: gap of unknown length

```


* 216918 217017: gap of unknown length
 * 217018 218872: contig of 1855 bp in length.

FEATURES

source 1..218872
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-210L21"
 1..2669
 /note="assembly_name:Contig100"
 2670..2769
 /estimated_length=unknown
 misc_feature 2770..4554
 /note="assembly_name:Contig101"
 gap 4555..4654
 /estimated_length=unknown
 misc_feature 4655..6908
 gap 6909..7008
 /note="assembly_name:Contig102"
 /estimated_length=unknown
 misc_feature 7009..10387
 /note="assembly_name:Contig103"
 gap 10388..10487
 /estimated_length=unknown
 misc_feature 10488..12825
 /note="assembly_name:Contig104"
 gap 12826..12925
 /estimated_length=unknown
 misc_feature 12926..15663
 /note="assembly_name:Contig105"
 gap 15664..15763
 /estimated_length=unknown
 misc_feature 15764..18992
 /note="assembly_name:Contig106"
 gap 18993..19092
 /estimated_length=unknown
 misc_feature 19093..21884
 /note="assembly_name:Contig107"
 gap 21885..21984
 /estimated_length=unknown
 misc_feature 21985..25307
 /note="assembly_name:Contig108"
 gap 25308..25407
 /estimated_length=unknown
 misc_feature 25408..38126
 /note="assembly_name:Contig109"
 gap 38127..38226
 /estimated_length=unknown
 misc_feature 38227..67816
 /note="assembly_name:Contig110"
 clone_end:T7
 vector_side:left
 gap 67817..67916
 /estimated_length=unknown
 misc_feature 67917..175443
 /note="assembly_name:Contig111"
 clone_end:SP6
 vector_side:right
 gap 175444..175543
 /estimated_length=unknown
 misc_feature 175544..176878
 /note="assembly_name:Contig142"
 gap 176879..176978
 /estimated_length=unknown
 misc_feature 176979..178260
 /note="assembly_name:Contig146"
 gap 178261..178360
 /estimated_length=unknown
 misc_feature 178361..179616
 /note="assembly_name:Contig151"
 gap 179617..179716
 /estimated_length=unknown

Query Match 91.6%; Score 17.4; DB 12; Length 218872;
 Best Local Similarity 94.7%; Pred. No. 3.6e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 123071 CATTACTTCATAGCTTTG 123053

RESULT 82

AC112307 234431 bp DNA linear HTG 13-NOV-2002
 Rattus norvegicus clone CH230-257P20, WORKING DRAFT SEQUENCE.
 AC112307
 AC112307.4 GI:24942081
 RTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 234431)
 Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Y., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C., Falte,L., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idiebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmood,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokemele,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Snajds,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,D., Yoon,V.,
 Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

Direct Submission

TITLE

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 234431)
Worley, K.C.
Direct Submission
Submitted (21-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 234431)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23603910.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMYQ
Center clone name: CH230-257P20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 173236 bases at least Q40
Consensus quality: 175680 bases at least Q30
Consensus quality: 176806 bases at least Q20
Estimated insert size: 179720; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 234431: contig of 234431 bp in length.
* Location/Qualifiers
  1. .234431
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-257P20"
  1. .1221
    /notes="wgs_end_extension
clone_end:T7"
46037. .47721
    /notes="wgs_end_extension
clone_end:T7"
54017. .54926
    /notes="clone_boundary
clone_end:T7"
site:
end sequence:B2162089"
56917. .57694
    /notes="clone_boundary

```

```

clone_end:Sp6
site:
end sequence:B2162091"
232904. .234431
/notes="wgs_end_extension
clone_end:Sp6"

misc_feature
1 CTTTACTTCATAGCTCTTTG 19
79129 CTTTATTTTCATAGCTCTTTG 79111

Query Match 91.6%; Score 17.4; DB 12; Length 234431;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

Rattus norvegicus clone CH230-84L6, WORKING DRAFT SEQUENCE.
AC096378 235763 bp DNA linear HTG 10-MAY-2003
Rattus norvegicus clone CH230-84L6, WORKING DRAFT SEQUENCE.
AC096378
AC096378.6 GI:30521763
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 235763)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguitano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chu, J.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Evans, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, W., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hoque, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Loulseg, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpungsoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sosa, J.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,

```

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 235763)
Worley, K. C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 235763)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24818898.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFAO
Center clone name: CH230-84L6
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 224184 bases at least Q40
Consensus quality: 226493 bases at least Q30
Consensus quality: 227839 bases at least Q20
Estimated insert size: 235519; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 235763: contig of 235763 bp in length.

FEATURES
source
Location/Qualifiers
1..235763
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-84L6"
misc_feature
1..1226
/note="wgs contig"
misc_feature
233871..235763
/note="wgs_contig"

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 235763;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATACCTCTTTG 19
|||||
Db 205307 CTTTACTTCATACCTTTG 205289

RESULT 84
AC095186
LOCUS
DEFINITION
Rattus norvegicus clone CH230-9B21, WORKING DRAFT SEQUENCE, 3 unordered pieces.
AC095186
AC095186.6 GI:24941004
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS
Muzny, D., Marie, H., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Fall, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, D., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulseged, H., Lozado, R., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milobavjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, D., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, B., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scheter, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C., Smajda, D., Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

TITLE

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 247690)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 247690)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT
On Nov 13, 2002 this sequence version replaced gi:2277986.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GDBT
Center clone name: CH230-9B21

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 215114 bases at least Q40
Consensus quality: 217886 bases at least Q30
Consensus quality: 219758 bases at least Q20
Estimated insert size: 221396; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 101341: contig of 101341 bp in length
* 101342 101441: gap of unknown length
* 101442 219685: contig of 118244 bp in length
* 219686 219785: gap of unknown length
* 219786 247690: contig of 27905 bp in length.
Location/Qualifiers
1. 247690
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9B21"
1. 1604
/notes="wgs contig"
101342 101441
/estimated_length=unknown
219686..219785
/estimated_length=unknown

FEATURES
source
misc_feature
gap
gap
ORIGIN
Query Match 91.6%; Score 17.4; DB 12; Length 247690;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;

```

```

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
|||||
Db 200451 CTTTACTTCAAAGCTTTG 200469

RESULT 85
AC137204/c
LOCUS AC137204
DEFINITION Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
***, 4 unordered pieces.
AC137204
VERSION AC137204.1 GI:25073092
HTG; HTGS PHAS1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 249669)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvai, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, W., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, C.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpethy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. I., Lebow, H., Levan, J., Lewis, D., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshehwa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C. D., Smajd, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlaczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., Smith
Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 249669)

```

AUTHORS
TITLE Rat Genome Sequencing Consortium.
JOURNAL Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KZPB
 Center clone name: CH230-unknown
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 201261 bases at least Q40
 Consensus quality: 206093 bases at least Q30
 Consensus quality: 208911 bases at least Q20
 Estimated insert size: 209892; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank draft data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 103590: contig of 103590 bp in length
 * 103591 103690: gap of unknown length
 * 103691 218904: contig of 115214 bp in length
 * 218905 219004: gap of unknown length
 * 219005 248412: contig of 29408 bp in length
 * 248413 248512: gap of unknown length
 * 248513 249669: contig of 1157 bp in length.

FEATURES
 source
 1. 249669
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-unknown"
 1. 1761
 /note="wgs contig"
 misc_feature
 4671..6761
 /note="wgs contig"
 gap
 103591..103690
 /estimated_length=unknown
 gap
 218905..219004
 /estimated_length=unknown
 gap
 248413..248512
 /estimated_length=unknown

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 249669;
 Best Local Similarity 94.7%; Pred. No. 3.6e+02;
 -Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTACTTCATAGCTTTG 19

Db 123053 CTTACTTCATAGCTTTG 123053

RESULT 86
 AC108958
 LOCUS
 DEFINITION

AC108958 265844 bp DNA linear HTG 13-MAY-2003
 Rattus norvegicus clone CH230-225123, *** SEQUENCING IN PROGRESS
 ***, 2 unordered pieces.

ACCESSION
 AC108958.5 GI:30579753
 VERSION
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 265844)

REFERENCE
 AUTHORS

Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensheva,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Mhlosoavjevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished
 2 (bases 1 to 265844)

Worley,K.C.

Direct Submission

Submitted (03-FEB-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 265844)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS TITLE JOURNAL

COMMENT

On May 13, 2003 this sequence version replaced gi:23321752.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GJEA
 Center clone name: CH230-225I23
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 241893 bases at least Q40
 Consensus quality: 247086 bases at least Q30
 Consensus quality: 249647 bases at least Q20
 Estimated insert size: 257860; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one 'clone'
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved

* 1 262982: contig of 262982 bp in length
 * 262983 263082: gap of unknown length
 * 263083 265844: contig of 2762 bp in length.

FEATURES

source

1. 265844
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-225I23"

misc_feature

1. 1246
 /note="wgs end extension
 clone_end:Sp6"

misc_feature

1870..3236
 /note="wgs end extension
 clone_end:Sp6"

misc_feature

3572..4405
 /note="clone boundary
 clone_end:Sp6
 site:EcORI

misc_feature

end sequence:BZ101420"
 complement(257901..258877)
 /note="clone boundary
 clone_end:T7
 site:EcORI
 end_sequence:BZ101420"

misc_feature 261286..262982
 /note="wgs end extension
 clone_end:T7"
 gap 262983..263082
 /estimated_length=unknown
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 12; Length 265844;
 Best Local Similarity 94.7%; Pred. No. 3.6e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGCTCTTG 19
 |||||
 Db 112412 CTTTACTTCATAGCTTTG 112430
 RESULT 87
 AC116211
 LOCUS Rattus norvegicus clone CH230-69K7, *** SEQUENCING IN PROGRESS ***,
 DEFINITION 4 unordered pieces.
 AC116211
 VERSION AC116211.4 GI:25007404
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 268030)
 Muzny,D., Marie., Metzker,M., Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyte,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,I., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., Lombard,H., Longacre,S., Lopez,J.,
 Lorensuewa,L., Loulseghe,H., Lozano,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokemeleh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojars,A., Rose,M., Rose,K., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,S.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 268030)
Worley, K.C.

Direct Submission
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 268030)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23269678.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center.
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GMYB
Center clone name: CH230-69K7
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 238815 bases at least Q40
Consensus quality: 241421 bases at least Q30
Consensus quality: 243001 bases at least Q20
Estimated insert size: 245676; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 241216: contig of 241216 bp in length
* 241217 241316: gap of unknown length
* 241317 264502: contig of 23186 bp in length
* 264503 264602: gap of unknown length
* 264603 266542: contig of 1940 bp in length
* 266543 266642: gap of unknown length
* 266643 268030: contig of 1388 bp in length.

FEATURES
source
Location/Qualifiers
1..268030
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-69K7"
1..1142

misc_feature
1..1142

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 241216: contig of 241216 bp in length
* 241217 241316: gap of unknown length
* 241317 264502: contig of 23186 bp in length
* 264503 264602: gap of unknown length
* 264603 266542: contig of 1940 bp in length
* 266543 266642: gap of unknown length
* 266643 268030: contig of 1388 bp in length.

FEATURES
source
Location/Qualifiers
1..268030
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-69K7"
1..1142

misc_feature
1..1142

misc_feature
1326..3278
/note="wgs contig"

gap
241217..241316
/estimated_length=unknown

gap
264503..264602
/estimated_length=unknown

gap
266543..266642
/estimated_length=unknown

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 268030;
Best Local Similarity 94.7%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
||||| ||||||| |||||||

Db 22540 CTTTACTTCATAGTCCTTTG 22558

RESULT 88

AB219542
LOCUS
DEFINITION
AB219542 4343 bp DNA linear BCT 02-JUL-2005
Campylobacter lari topA, flaA, flaB, CLA0521 genes for
Topoisomerase I, Flagellins, hypothetical protein CLA0521, partial
and complete cds, strain: NCTC12892.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AB219542.1 GI:68533170
Campylobacter lari
Campylobacter lari
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 4343)
Sekizuka, T., Murayama, O., Moore, J.E., Millar, B.C. and Matsuda, M.
Variations of two flagellin genes of Campylobacter lari, in
particular urease-positive thermophilic Campylobacter (UPTC)
Unpublished

2 (bases 1 to 4343)
Sekizuka, T., Murayama, O., Moore, J.E., Millar, B.C. and Matsuda, M.
Submitted (30-JUN-2005) Motoo Matsuda, Azabu University, Laboratory
of Molecular Biology, School of Environmental Health Sciences;
1-17-71, Fuchinobe, Kanagawa, 229-8501, Japan
(E-mail: matsuda@azabu-u.ac.jp, tel: 81-42-769-1942,
Fax: 81-42-754-7661)
Location/Qualifiers
1..4343
/organism="Campylobacter lari"
/mol_type="genomic DNA"
/strain="NCTC12892"
/isolation_source="River water"
/db_xref="taxon:201"
/note="Urease-positive thermophilic Campylobacter"
complement(<1..874)
/gene="topA"
complement(<1..822)
/gene="topA"
/codon_start=1
/transl_table=11
/product="Topoisomerase I"
/protein_id="BA806130.1"
/db_xref="GI:68533171"
/translation="MKNLIIIVSPAKAKTIGNFGDKDYEVIAKSHIRDLPKTSFGIK
IEDENPKFYRISNDHSLVKEKAKAKTIYLADEDEGEAIAYHAKAINKDE
NSLPRIVFHEITKSAIENALNPRSLNNSVNAQOTRRLLDRIYGYKLSPLNKKIOK
GLSAGRVQSAALKIIVDREIEIKAFVPLKYSIDIMFEKDLQAEVLFQNKIKELSL
TNEERAKLIFEACNANFNKOTESKDKRIAPQAPFTWTSLQOSASRLGNPKKT"
complement(828..831)
/gene="topA"
/note="putative RBS"
complement(843..848)
/gene="topA"

gene
CDS
RBS
-10_signal


```
/note="putative prpbnow box"
complement(868. .874)
/genes="copA"
/notes="putative -35 region"
complement(887. .2477)
/genes="flaA"
complement(887. .918)
/genes="flaA"
/notes="putative transcriptional terminator"
complement(936. .2399)
/genes="flaA"
/codon_start=1
/transl_table=11
/product="Flagellin"
/protein_id="BAE06131.1"
/db_xref="GI:68533172"
/translation="MGPRINTNVASLNAOVNAGNSPALDSSLRLSSGLRINSAAADD
ASGVAIADSLRNQAASLGQALNNNDRAIGLQADKAMDEQLKILDTIKVKATQAAQD
GQSAKTRAMIQGEINKLMEELDNIAFTYNGKQLLSGAFNQOQFQVGDKANOTINAT
IGATQSAKIQGTFRFETGSRITGSGNAGFTIKNYDGVNDFKIQSVILSTAGTGLGALA
AEINKSSDKTGVTRATVQTISSGTIQAGNTGDTFTINGVIGKVAQADKDGSLVA
AINAKDPTTGVASVNGQVLNSADGRGIELSGLTALSGNIAVNVYGRSLVKNQD
SDIILSGSGAGRGTAARAEATVNLSEVKGOISADIACAMGNASATOPGSKTGCVT
TLQAMVMDIADTAIANDTIRANIGATONQITSTINNISVTQNVKAAESQIRDVD
FASANSYKANILAQSGSYAMAQANAASQNVLRLLQ"
/genes="flaA"
complement(2406. .2409)
/genes="flaA"
/notes="putative RBS"
complement(2450. .2455)
/genes="flaA"
/notes="putative prpbnow box"
complement(2472. .2477)
/genes="flaA"
/notes="putative -35 region"
2508. .4037
/genes="flaB"
2508. .2509
/genes="flaB"
/notes="putative -26 signal"
2520. .2521
/genes="flaB"
/notes="putative -13 signal"
2549. .2552
/genes="flaB"
/notes="putative RBS"
2559. .4037
/genes="flaB"
/codon_start=1
/transl_table=11
/product="Flagellin"
/protein_id="BAE06132.1"
/db_xref="GI:68533173"
/translation="MGPRINTNIGAMNAHANSVITARELKSIGRLSSGLRINSAAADD
ASGVAIADSLRNQAASLGQALNNNDRAIGLQADKAMDEQLKILDTIKVKATQAAQD
GQSAKTRAMIQGEINKLMEELDNIAFTYNGKQLLSGAFNQOQFQVGDKANOTINAT
IGATQSAKIQGTFRFETGSRITGSGNAGFTIKNYDGVNDFKIQSVILSTAGTGLGALA
AEINKSSDKTGVTRATVQTISSGTIQAGNTGDTFTINGVIGKVAQADKDGSLVA
AINAKDPTTGVASVNGQVLNSADGRGIELSGLTALSGNIAVNVYGRSLVKNQD
SDIILSGSGAGRGTAARAEATVNLSEVKGOISADIACAMGNASATOPGSKTGCVT
TLQAMVMDIADTAIANDTIRANIGATONQITSTINNISVTQNVKAAESQIRDVD
FASANSYKANILAQSGSYAMAQANAASQNVLRLLQ"
complement(4044. .>4343)
/genes="CLA0521"
complement(4044. .>4343)
/genes="CLA0521"
/codon_start=1
/transl_table=11
/product="hypothetical protein CLA0521"
/protein_id="BAE06133.1"
/db_xref="GI:68533174"
/translation="DLTQDIDKFOILLIKIDNFKVSYNNKQPFYFTEISKPLFYHVE
CKISELFIKNNDKADKKKKSVEYLNIETKLWINDYILNLIKQNDLLKQTPLYQGV"
4045. .4074
```

```
terminator
/genes="flaB"
4045. .4074
/genes="flaB"
/notes="putative transcriptional terminator"

ORIGIN
Query Match 89.5%; Score 17; DB 15; Length 4343;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
Db 744 TTACTTCATAGTCTTTG 760
|||||

RESULT 89
AC172882
LOCUS AC172882 67767 bp DNA linear HTG 24-NOV-2005
DEFINITION Brassica rapa subsp. pekinensis clone KBrH121H08, *** SEQUENCING IN
PROGRESS ***, 2 ordered pieces.
ACCESSION AC172882
VERSION AC172882.1 GI:82654463
KEYWORDS HTG; HTGS PHASE2
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 67767)
AUTHORS Yang,T.J., Kwon,S.J., Kim,J.A., Park,J.Y., Kim,J.S., Jin,M.,
Lim,K.B., Lim,M.H., Choi,B.S., Lim,Y.P. and Park,B.S.
CONSTRM Korea Brassica Genome Project (KBGP)
TITLE Shotgun sequence of a Brassica rapa BAC clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 67767)
AUTHORS Yang,T.J., Kwon,S.J., Kim,J.A., Park,J.Y., Kim,J.S., Jin,M.,
Lim,K.B., Lim,M.H., Choi,B.S., Lim,Y.P. and Park,B.S.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-2005) Brassica Genomics Team, National Institute
of Agricultural Biotechnology (NIAB), RDA, 224 Seodun-dong, Suwon,
Kyung-Ki Do 441-707, Korea (E-mail:pbeom@rda.go.kr;
Tel:82-31-239-1670; Fax:82-31-299-1672)
COMMENT * NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 12938: contig of 12938 bp in length
* 12939 13038: gap of unknown length
* 13039 67767: contig of 54729 bp in length.
FEATURES
source
1. 67767
/organism="Brassica rapa subsp. pekinensis"
/mol_type="genomic DNA"
/cultivar="Inbred line 'Chiifu'"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/clone="KBrH121H08"
/tissue_type="Young Leaves"
/clone_lib="KBrH (HindIII) BAC Library"
12939_13038
/estimated_length=unknown
gap
12939_13038
/estimated_length=unknown

ORIGIN
Query Match 89.5%; Score 17; DB 12; Length 67767;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
```

```
Db      15936 TTTACTTCATAGTCTTT 15952
|||||
RESULT 90
CP000082_03/c
WPCOMMENT
Sequence split into 27 fragments LOCUS CP000082 Accession CP000082
Fragment Name      Begin      End
CP000082_00       1      110000
CP000082_01      100001   210000
CP000082_02      200001   310000
CP000082_03      300001   410000
CP000082_04      400001   510000
CP000082_05      500001   610000
CP000082_06      600001   710000
CP000082_07      700001   810000
CP000082_08      800001   910000
CP000082_09      900001  1010000
CP000082_10     1000001  1100000
CP000082_11     1100001  1210000
CP000082_12     1200001  1310000
CP000082_13     1300001  1400000
CP000082_14     1400001  1510000
CP000082_15     1500001  1610000
CP000082_16     1600001  1710000
CP000082_17     1700001  1810000
CP000082_18     1800001  1910000
CP000082_19     1900001  2010000
CP000082_20     2000001  2110000
CP000082_21     2100001  2210000
CP000082_22     2200001  2310000
CP000082_23     2300001  2410000
CP000082_24     2400001  2510000
CP000082_25     2500001  2610000
CP000082_26     2600001  2650701
Continuation (4 of 27) of CP000082 from base 300001 (CP000082 Psychrobacter arcticus 273

Query Match      89.5%; Score 17; DB 15; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
|||||
Db      63421 TTTACTTCATAGTCTTT 63405
|||||

RESULT 91
AE008384_01
WPCOMMENT
Sequence split into 41 fragments LOCUS AE008384 Accession AE008384
Fragment Name      Begin      End
AE008384_00       1      110000
AE008384_01      100001   210000
AE008384_02      200001   310000
AE008384_03      300001   410000
AE008384_04      400001   510000
AE008384_05      500001   610000
AE008384_06      600001   710000
AE008384_07      700001   810000
AE008384_08      800001   910000
AE008384_09      900001  1010000
AE008384_10     1000001  1110000
AE008384_11     1100001  1210000
AE008384_12     1200001  1310000
AE008384_13     1300001  1410000
AE008384_14     1400001  1510000
AE008384_15     1500001  1610000
AE008384_16     1600001  1710000
AE008384_17     1700001  1810000
AE008384_18     1800001  1910000
AE008384_19     1900001  2010000
AE008384_20     2000001  2110000
```

```
AE008384_21      2100001  2210000
AE008384_22      2200001  2310000
AE008384_23      2300001  2410000
AE008384_24      2400001  2510000
AE008384_25      2500001  2610000
AE008384_26      2600001  2710000
AE008384_27      2700001  2810000
AE008384_28      2800001  2910000
AE008384_29      2900001  3010000
AE008384_30      3000001  3110000
AE008384_31      3100001  3210000
AE008384_32      3200001  3310000
AE008384_33      3300001  3410000
AE008384_34      3400001  3510000
AE008384_35      3500001  3610000
AE008384_36      3600001  3710000
AE008384_37      3700001  3810000
AE008384_38      3800001  3910000
AE008384_39      3900001  4010000
AE008384_40      4000001  4096345
Continuation (2 of 41) of AE008384 from base 100001 (AE008384 Methanosarcina mazei strain

Query Match      89.5%; Score 17; DB 15; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
|||||
Db      73057 TTTACTTCATAGTCTTT 73073
|||||

RESULT 92
AE010299_40
WPCOMMENT
Sequence split into 58 fragments LOCUS AE010299 Accession AE010299
Fragment Name      Begin      End
AE010299_00       1      110000
AE010299_01      100001   210000
AE010299_02      200001   310000
AE010299_03      300001   410000
AE010299_04      400001   510000
AE010299_05      500001   610000
AE010299_06      600001   710000
AE010299_07      700001   810000
AE010299_08      800001   910000
AE010299_09      900001  1010000
AE010299_10     1000001  1110000
AE010299_11     1100001  1210000
AE010299_12     1200001  1310000
AE010299_13     1300001  1410000
AE010299_14     1400001  1510000
AE010299_15     1500001  1610000
AE010299_16     1600001  1710000
AE010299_17     1700001  1810000
AE010299_18     1800001  1910000
AE010299_19     1900001  2010000
AE010299_20     2000001  2110000
AE010299_21     2100001  2210000
AE010299_22     2200001  2310000
AE010299_23     2300001  2410000
AE010299_24     2400001  2510000
AE010299_25     2500001  2610000
AE010299_26     2600001  2710000
AE010299_27     2700001  2810000
AE010299_28     2800001  2910000
AE010299_29     2900001  3010000
AE010299_30     3000001  3110000
AE010299_31     3100001  3210000
AE010299_32     3200001  3310000
AE010299_33     3300001  3410000
AE010299_34     3400001  3510000
AE010299_35     3500001  3610000
AE010299_36     3600001  3710000
```

```
AE010299_37 3700001 3810000
AE010299_38 3800001 3910000
AE010299_39 3900001 4010000
AE010299_40 4000001 4110000
AE010299_41 4100001 4210000
AE010299_42 4200001 4310000
AE010299_43 4300001 4410000
AE010299_44 4400001 4510000
AE010299_45 4500001 4610000
AE010299_46 4600001 4710000
AE010299_47 4700001 4810000
AE010299_48 4800001 4910000
AE010299_49 4900001 5010000
AE010299_50 5000001 5110000
AE010299_51 5100001 5210000
AE010299_52 5200001 5310000
AE010299_53 5300001 5410000
AE010299_54 5400001 5510000
AE010299_55 5500001 5610000
AE010299_56 5600001 5710000
AE010299_57 5700001 5751492
Continuation (41 of 58) of AE010299 from base 4000001 (AE010299 Methanosarcina acetivorans)

Query Match 89.58; Score 17; DB 15; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCCTT 18
Db 76664 TTTACTTCATAGTCCTT 76680

RESULT 93
AC155344 110038 bp DNA linear PLN 12-AUG-2005
LOCUS Brassica rapa subsp. pekinensis clone KBrH080A08, complete
DEFINITION
AC155344
AC155344.1 GI:57900806
VERSION
KEYWORDS HTG.
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 110038)
Yang T.J., Kim J.S., Kwon S.J., Kim J.A., Lim K.B., Jin M.,
Park J.Y., Lim M.H., Lim Y.P., Choi B.S. and Park B.S.
KEGG
Complete sequence of a Brassica rapa BAC clone
Unpublished
2 (bases 1 to 110038)
Yang T.J., Kim J.S., Kwon S.J., Kim J.A., Lim K.B., Jin M.,
Park J.Y., Lim M.H., Lim Y.P. and Park B.S.
Direct Submission
Submitted (19-JAN-2005) Brassica Genomics Team, National Institute
of Agricultural Biotechnology (NIAB), RDA, 224 Seodun-dong, Suwon,
Kyunggi-Do 441-707, Korea (E-mail: pbcom@rda.go.kr;
Tel: 82-31-299-1670; Fax: 82-31-299-1672)
3 (bases 1 to 110038)
Yang T.J., Kim J.S., Kwon S.J., Kim J.A., Lim K.B., Jin M.,
Park J.Y., Lim M.H., Lim Y.P. and Park B.S.
Direct Submission
Submitted (12-AUG-2005) Brassica Genomics Team, National Institute
of Agricultural Biotechnology (NIAB), RDA, 224 Seodun-dong, Suwon,
Kyunggi-Do 441-707, Korea (E-mail: pbcom@rda.go.kr;
Tel: 82-31-299-1670; Fax: 82-31-299-1672)
This is a finished sequence consisting of 1 contig high quality.
The BAC clone contains flowering locus C gene and located on the
cytogenetic chromosome 10 by genetic map and FISH analyses.
FEATURES
source
1..110038
/organism=Brassica rapa subsp. pekinensis"
```

```
/mol_type="genomic DNA"
/cultivar="Inbred line 'Chiifu'"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/clone="KBrH080A08"
/tissue_type="Young Leaves"
/clone_lib="KBrH (HindIII) BAC Library"
17..2486
/genes="80A08_1"
/notes="Homolog of At5g10020, receptor-like protein kinase
ERECTA, Arabidopsis thaliana, EMBL:AC004484; go_function:
protein serine/threonine kinase activity [goid 0004674];
go_function: ATP binding [goid 0005524]; go_process:
protein amino acid phosphorylation [goid 0006468];
go_process: transmembrane receptor protein tyrosine kinase
signaling pathway [goid 0007169]"
join(17..1256,1332..1909)
/genes="80A08_1"
/notes="Homolog of At5g10020"
/codon_start=1
/product="80A08_1"
/protein_id="AAZ67586.1"
/db_xref="GI:72384470"
/translation="MNSLEGPIPSGRSASELMAISSDPOMELDLSTNLTGALPGD
IGTMEIRIVLVNANKLSGELPSELNKLTSLEFLDLSNDFKQIPEKPSGWARFNV
SYNDLSGTIPENLRSYPHSFYGNKLRIPGTPLGSSNGVLAVQKADHHSRLSI
KTAIVASGAAALMVLVLFVYHRTQKDFHVNQATTRDAKGRSRPFLNPTSNA
EQOSSLSFSNDHLTFANSRSLSGIPGSEPIEQALPAASATTINLLDDYPATSGR
KSTSGGSLSPRFPQVMDLVYSPDLRAGELFFLDVSLKLTAEELSRAPAYVWGR
SGTLYKATLDNGHMLTVKWLVLGVRHKDKFAKKIKGLKHPNIVPURYVWGR
EQERLLISDYMGESLAMLHYETTPRYPSPMSFIQRLKVAVEVAQCLYLHDRAMPHG
NLKPTNIILTPENTVRIIDYCIHRLSSGVAEQILNMSALGYSAPELASAKPVPT
LKSDVYAFGVILMELLTRRSAGDIISGQSGAVLTDVRLCDQEGRRMDCIDRDIAGG
EEPSKTMEDALAIKCIASVNERPIRQVLDLILSLSS"
/genes="80A08_2"
/notes="Homolog of At5g10030, identical to ocs-element
binding factor 4 GI:414613 from [Arabidopsis thaliana];
go_function: DNA binding [goid 0003677]; go_function:
transcription factor activity [goid 0003700]; go_function:
calmodulin binding [goid 0005516]"
complement(join(5074..5226,5298..5525,5610..5881,
5957..6020,6094..6171,6267..6326,6400..6471,6547..6714))
/genes="80A08_2"
/notes="Homolog of At5g10030"
/codon_start=1
/product="80A08_2"
/protein_id="AAZ67587.1"
/db_xref="GI:72384471"
/translation="MNTTSTHFVPTTRFEIYDPLNQTGWEEESFKNNGGGFTPN
IITPTNQKYSLEDEGTGTPHFDQEAESTSRHPDKTRQLAQNREAAKSRURKKAY
VQOLETSLRLKLIHLEQELDRARQQGYFASNRVDTNALSFDNMCSGIVAFEMEGHW
EQNRROISELRTVLNGQVSDIELRLVDMNKHYFOLFRMKSAALKLDVFIYINSQMKL
TSAERFLATGGFRPSSELLKVLPHFDDMMQDQVLDVCLNRQSCQQAEDAVSQMEKL
QHTLAESVAGELGEGSYVPQITSAMERLEALVSFVNQADHLRHETLQOMHRLITTRQ
AARGLALGVEYFQRLKALSSSWETROREPT"
10880..11395
/genes="80A08_3"
/notes="Homolog of At5g10040, synonym: T31P16.30; expressed
protein"
10954..11178
/genes="80A08_3"
/notes="Homolog of At5g10040"
/codon_start=1
/product="80A08_3"
/protein_id="AAZ67588.1"
/db_xref="GI:72384472"
/translation="MADIALLVVEEYERRUKLTSEKASNSABTGWNDFFARVKMFGS
DQKRMESLKRKLKSAQFALVNSGFFSA"
complement(11543..14485)
/genes="80A08_4"
/notes="Homolog of At5g10070, contains Pfam profiles
PF04034: Domain of unknown function (DUF367), PF04068:
gene
CDS
gene
CDS
gene
CDS
gene
CDS
gene
CDS
gene
```

CDS
Possible metal-binding domain in RNase L inhibitor, RLI"
complement(join(11739..12038,12258..12420,12663)..12870,
12974..13089,13183..13254,13346..13436,13527..13663,
13755..13795,13893..13958,14063..14109,14193..14221,
14307..14386))
/gene="80A08_4"
/note="Homolog of At5g10070"
/codon_start=1
/product="80A08_4"
/protein_id="AAZ67589.1"
/db_xref="GI:72384473"
/translation="MGYGRHRRSGNSNPQGTERTENIGREDSLPHDPGSEDAVV
PNVLAMWDFQCQDKTRGLARFNLLKELRVNTGGVVLSPVQCVCSEDAV
IKSLGAVDCSWARLTDVPAKRLPMLVAANPVNPGRCVCEALSAA
LILGEBETANLLGKFKMGHAFSLNKLIIKEYSKCNSAEIISVONSWLTOQTLS
KQAPLKHVRKGEDEDDDDGLPPLERMMHVISDESEDEDEDEDEDDGLPPL
ERNVHVLGDEEDDSEGLCLCTSSCGVAEKATASAYVQNVNNESESEKCK
SAVKRIMLKKPALLTLVQGHVIANPAAPQFLKPPVNPYFGNIFLMPGPP
PHMIGNQPNQPSNSAQGPALQPPGMIYYRAPHRGSSGCTILFLVFLTYPALFH"
14643..18263
/gene="80A08_5"
/note="Homolog of At5g10080, contains Eukaryotic and viral
aspartyl proteases active site, PROSIRE:PS00141;
go_function: peptin A activity [gold 0004194]; go_process:
proteolysis and peptidolysis [gold 0006508]"
join(14926..15209,15305..15443,15583..15821,15902..16129,
16210..16419,16495..16588,16680..16743,16843..16914,
17004..17197,17297..17523,17967..18114)
/gene="80A08_5"
/note="Homolog of At5g10080"
/codon_start=1
/product="80A08_5"
/protein_id="AAZ67590.1"
/db_xref="GI:72384474"
/translation="MAKRSFILLFILSLVSEKSLASLFSSRLIHRFSDEGRASIKSP
GSFPERSFLEYRLLTIDRRQNLGAKFQSLVPSEGSTISFGNYFGWLHWTWID
IGTSPSVFLVALDGSLLMTPCNVCQAPLSSAYYSLATKOLNEPDPASASTSKVF
PCSHLCESAPACSPKEQCPYTYTASENTSSGLAVEDVLHAYSANASSVVKARV
VVGCGKSGEFLKGIAPDVGMLGRGEISVPSFLAKGLMNSFCFDEEDSGRIY
FGDVPSTQSTRFLPNYNEFVAFGVEVCCVNSCLQSSFTILDSGGSTFHLPLP
EYREVALEIDSHINATVKIEGPEYCYETSPEPKVPAIKLFESNNFTVFIHKPLF
VLORSEGLVDFCLPISASEEGTGVIQONYMAGRIYVFDRENKLGHSASKQCDKIA
PQPASPGSTSNPNLPTEQOQTHAVSPAIGKTPSKTSSACCSFNRLISDLSIL
LGKPAQASTSAHAPRDTCRAGEGTIFFRVKLWFYASHDLGLGNAIAIRVRLGFEI
AIFALFSSLSLSLSTHGAVKVEVTFMAEEDWIVKLDPAQFEKVADSVELLCP
GLGRCDGG"
complement(19424..22137)
/gene="80A08_6"
/note="hypothetical protein"
complement(join(20639..20875,20911..21065,21363..21594))
/gene="80A08_6"
/note="hypothetical protein"
/codon_start=1
/product="80A08_6"
/protein_id="AAZ67591.1"
/db_xref="GI:72384475"
/translation="MQPFLSKLENKPEDRLPLVYNQATQSSQATQSCQATLNPPSOS
SQATQSNPATQSSASQSPATKNPADVTSSTWKRLGSCASRGPIPIYPLVIT
RVCNVSSTGYLVIIFLNGVKVYINTLPEVPTPEVREDEILDLPDQVRLSEL
LHLILALIRQLQRFLLKGDPMYDSNGYRERDFILVISLTYIS"
complement(22215..25674)
/gene="80A08_7"
/note="Homolog of At5g10090, contains Pfam profile PF00515
TPR Domain"
complement(join(22472..22606,22684..23017,23110..23478,
23557..24401))
/gene="80A08_7"
/note="Homolog of At5g10090"
/codon_start=1
/product="80A08_7"
/protein_id="AAZ67592.1"
/db_xref="GI:72384476"
/translation="MAENQAEKRSGRLGLGFVFGRRGLWSKKCTADNGNKTTPMRSSNA

gene
SAPCTSNIQFTKSPGNELNSKKLOEYKVSPEPIQONOTOIORPISKPLSNQYPNNNN
GPVQOARKVVRESIGLSGLESMTDNOKAKGMFCNLGNLKQPGTTAVGNQTTVO
NSGVRKTMERGEOTPYRPTSVSNQOQSSGLCSRAIETMDPEILKINGEDYKNGNF
VEALYDAAIAIDPKAAKSNKSAALTALGRIDEAVFECEAIRMEPHYRAHRL
ANLYLRIGEVENSIIHKRSGPEADQDILKAKTVQHLNCKTEAKRLDWNLLKET
KNTIAGDAATQVYALQAEAFKSYRHEADDALSCPCFVDFVENKTKYIGYAGF
LVVWAQVHMSLGRGEAVEATQLAAKLDNRNREVMVLRVQVATTAARSKGNDFKCTG
RFQASAAVSGRGLDHSRNSVLCNRAACLFKMGQFDRAGIDGTAALSVRPAYAKARL
RRADCNAKLGNWELAVGDYELRKETPEDDDQVIRGLMEAQHLYKRGHENL"
28210..32034
/gene="80A08_8"
/note="Homolog of At5g10100, similar to
trehalose-6-phosphate phosphatase (AtPFPB) [Arabidopsis
thaliana] GI:2944180; contains Pfam profile PF02358:
Trehalose-phosphatase; go_function: trehalose-phosphatase
Query Match 89.5%; Score 17; DB 4; Length 110038;
Best Local Similarity 100.0%; Pred. No. 5.6e+07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 TTTACTTCATAGTCTTT 18
|||||
Db 7684 TTTACTTCATAGTCTTT 7700
RESULT 94
AL359642/c
LOCUS AL359642 122613 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-507F17 on chromosome 9 Contains
a kertain pseudogene, complete sequence.
ACCESSION AL359642
VERSION AL359642.12 GI:16972925
KEYWORDS HTG; keratin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 122613)
Dunn, M.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegasanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 16, 2001 this sequence version replaced gi:15591063.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-507F17 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegasanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1. .122613
FEATURES:
source

Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrocino, J., Pham, T., Primus, E., Pu, L.-L., Puzo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rivers, C., Rodriguez, F., Rojas, A., Ruiz, S.-J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatok, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villaseña, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

Unpublished
Direct Submission
2 (bases 1 to 145877)
Worley, K.C.
Direct Submission
Submitted (20-JAN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 145877)
Worley, K.C.
Direct Submission
Submitted (27-MAY-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 145877)
Worley, K.C.
Direct Submission
Submitted (29-MAY-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 29, 2005 this sequence version replaced gi:66773448.
Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Drafting Center Code: WIBR
Contact: hgsc-help@bcm.tmc.edu.
Location/Qualifiers
1. 145877
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP24-399D24"
complement(41..188)
/rpt_family="B2 Mm2"
complement(170..263)
/rpt_family="B1 Mur3"
complement(211..281)
/rpt_family="B4"
282..333
/rpt_family="AT rich"
complement(359..1324)
/rpt_family="RMER178"
complement(1157..1821)
/rpt_family="LX9"
1823..1842
/rpt_family="(T)n"

```
repeat_region complement(1843..1934)
repeat_region /rpt family="B1_Mus2"
repeat_region complement(1934..1998)
repeat_region /rpt family="MTD"
repeat_region complement(2013..2685)
repeat_region /rpt family="Lx4B"
repeat_region complement(2686..2967)
repeat_region /rpt family="Lx5"
repeat_region 3366..3475
repeat_region /rpt family="(TCTA)n"
repeat_region complement(3477..3556)
repeat_region /rpt family="B1_Mus1"
repeat_region 3577..3629
repeat_region /rpt family="(CA)n"
repeat_region 3676..3697
repeat_region /rpt family="AT_rich"
repeat_region 3733..3920
repeat_region /rpt family="B3"
STS 4490..4771
/standard_name="RH125353"
repeat_region 5038..5066
repeat_region /rpt family="(CAAAA)n"
repeat_region 5080..5206
repeat_region /rpt family="B1_Mur1"
repeat_region 5605..5658
repeat_region /rpt family="(CCA)n"
repeat_region 5809..5942
repeat_region /rpt family="B3A"
repeat_region 5966..6129
repeat_region /rpt family="ID_B1"
repeat_region 5996..6130
repeat_region /rpt family="B1_Mur4"
repeat_region 6215..6333
repeat_region /rpt family="RSINE1"
repeat_region complement(6623..6682)
repeat_region /rpt family="B3A"
repeat_region complement(7735..7793)
repeat_region /rpt family="MurSatRep1"
repeat_region 8278..8412
repeat_region /rpt family="B1_Mur4"
repeat_region 8418..8446
repeat_region /rpt family="(A)n"
repeat_region complement(8503..8555)
repeat_region /rpt family="B2_Mm2"
repeat_region complement(8573..8744)
repeat_region /rpt family="B2_Mm1"
repeat_region complement(9439..9550)
repeat_region /rpt family="PB1D9"
repeat_region complement(9577..9696)
repeat_region /rpt family="ID_B1"
repeat_region 9724..9913
repeat_region /rpt family="B3"
repeat_region complement(10256..10404)
repeat_region /rpt family="RSINE1"
repeat_region 10496..10829
repeat_region /rpt family="RMER6A"
repeat_region 11114..11309
repeat_region /rpt family="B3"
repeat_region 11344..11535
repeat_region /rpt family="B2_Mm2"
repeat_region 11686..11751
repeat_region /rpt family="B3A"
repeat_region complement(11821..12010)
repeat_region /rpt family="B3A"
repeat_region complement(12016..12160)
repeat_region 13278..13330
repeat_region /rpt family="AT_rich"
repeat_region 14631..14740
repeat_region /rpt family="B3"
repeat_region 14795..14942
repeat_region /rpt family="B1_Mus1"
repeat_region complement(15333..15395)

/rpt family="RSINE1"
15644..15665
/rpt family="AT_rich"
complement(16347..16538)
/rpt family="B2_Mm1"
16962..17024
/rpt family="ID"
complement(18114..18158)
/rpt family="PB1D9"
complement(18222..18330)
/rpt family="B3A"
complement(18374..18464)
/rpt family="B3A"
complement(18600..18706)

Query Match 89.5%; Score 17; DB 6; Length 145877;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
|||||
Db 22360 TTTACTTCATAGTCTTT 22376

RESULT 96
CR450699/c
LOCUS CR450699.3 182422 bp DNA linear HTG 12-JAN-2006
DEFINITION Danio rerio chromosome 25 clone DKEIP-120A11, WORKING DRAFT
SEQUENCE CR450699
ACCESSION CR450699.3 GI:84871856
VERSION HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Danio rerio (zebrafish)
SOURCE ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 182422)
AUTHORS McLaren,S.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-2006) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
On Jan 12, 2006 this sequence version replaced gi:67508928.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: zkpi20A11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 180066 bases at least Q40
Consensus quality: 180649 bases at least Q30
Consensus quality: 181091 bases at least Q20
Insert size: 181722; sum-of-contigs
Insert size: 189182; 4.3% error; agarose-fp
Quality coverage: 6.23x in Q20 bases; sum-of-contigs Quality
coverage: 6.11x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 11550: contig of 11550 bp in length
* 11551 11650: gap of 100 bp
```

```

* 11651 26415: contig of 14765 bp in length
* 26416 26515: gap of 100 bp
* 39652 39653: contig of 13138 bp in length
* 39654 39753: gap of 100 bp
* 39754 42101: contig of 2348 bp in length
* 42102 42201: gap of 100 bp
* 42202 158128: contig of 115927 bp in length
* 158129 158228: gap of 100 bp
* 158229 163016: contig of 4788 bp in length
* 163017 163116: gap of 100 bp
* 163117 166652: contig of 3536 bp in length
* 16653 166752: gap of 100 bp
* 166753 182422: contig of 15670 bp in length.

FEATURES
    source
        1..182422
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /chromosome="25"
            /clone_lib="DanioKeyPilot"
            /clone_lib="DanioKeyPilot"
        1..11550
            /note="assembly fragment:00170"
            /fragment_chain:1
        11651..26415
            /note="assembly fragment:00086"
            /fragment_chain:1
        26516..39653
            /note="assembly fragment:00415"
            /fragment_chain:1
        39754..42101
            /note="assembly fragment:00011"
            /fragment_chain:2
        42202..158128
            /note="assembly fragment:02027"
            /fragment_chain:2
        158229..163016
            /note="assembly fragment:00018"
            /fragment_chain:1
        163117..166652
            /note="assembly fragment:00046"
            /fragment_chain:1
        166753..182422
            /note="assembly fragment:00270.0"

ORIGIN
    Query Match      89.5%; Score 17; DB 12; Length 182422;
    Best Local Similarity 100.0%; Pred. No. 5.8e+02;
    Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
Db 18233 TTTACTTCATAGTCTTT 18217

RESULT 97
AC148203/c
LOCUS AC148203 Callicebus moloch clone LB5-190B16, complete sequence.
DEFINITION AC148203
ACCESSION AC148203
VERSION AC148203.3 GI:67078611
KEYWORDS HTG
SOURCE Callicebus moloch (Dusky titi)
ORGANISM Callicebus moloch
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
Cebidae; Callicebinae; Callicebus.
1 (bases 1 to 187401)
Attonellis,A., Ayele,K., Bass,D., Benjamin,B., Bera,J.,
Blakesley,R.W., Bouffard,G., Brinkley,C., Brooks,S., Chu,G.,
Coleman,H., Engle,J., Fukeenko,T., Gestole,M., Greene,A., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hunter,G., Hurler,B., Idol,J.R., Kwong,P., Latic,P., Larson,S.,
Lee-tin,S.-O., Legaspi,R., Madden,M., Maduro,Q.I., Maduro,V.B.,
Margulies,E.H., Masello,C., Maskeri,B., McDowell,J., Mojidi,H.A.,
Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
Puri,O., Reddix-Dugue,N., Sante,A., Schandler,K., Schueler,M.G.,
Simon,C., Stantripop,S., Stephen,E., Taye,A., Thomas,J.W.,
Thomas,P., Tsipouri,V., Ung,L., Vogt,J.L., Wetherby,K.D.,
Withers,T.R., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 187401)
Green,E.D.
Direct Submission
Submitted (12-FEB-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 187401)
Green,E.D.
Direct Submission
Submitted (04-MAR-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
4 (bases 1 to 187401)
Green,E.D.
Direct Submission
Submitted (09-JUN-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Jun 9, 2005 this sequence version replaced gi:44917650.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: gte
Center clone name: 190B16

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

FEATURES
    source
        1..187401
            /organism="Callicebus moloch"
            /mol_type="genomic DNA"
            /db_xref="taxon:9523"
            /clone_lib="LB5-190B16"
            /clone_lib="LB5"
            /note="BAC resource: http://bacpac.chori.org/"
        48920..48927
            /note="single clone coverage"
        127297..127302
            /note="low quality single stranded/single chemistry region"
        135566..135658
            /note="single clone coverage"

ORIGIN
    Query Match      89.5%; Score 17; DB 5; Length 187401;
    Best Local Similarity 100.0%; Pred. No. 5.8e+02;
    Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
Db 71590 TTTACTTCATAGTCTTT 71574

RESULT 98
CR354430

```


LOCUS CR354430 195080 bp DNA linear HTG 08-FEB-2006
DEFINITION Danio rerio chromosome 25 clone CH211-218P2, WORKING DRAFT
SEQUENCE, 2 unordered pieces.
ACCESSION CR354430
VERSION CR354430.13 GI:87080614
KEYWORDS HTG; HTGS PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 195080)
AUTHORS Heath, P.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2006) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 zfish-help@sanger.ac.uk Clone requests:
 http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
 On Feb 8, 2006 this sequence version replaced gi:84616971.
COMMENT ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: zfish-help@sanger.ac.uk
 ----- Project Information
 Center project name: zc218P2
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 194916 bases at least Q40
 Consensus quality: 194943 bases at least Q30
 Consensus quality: 194965 bases at least Q20
 Insert size: 194980; sum-of-contigs
 Insert size: 181200; 14.5% error; agarose-fp
 Quality coverage: 12.17x in Q20 bases; sum-of-contigs Quality
 coverage: 13.28x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 106591: contig of 106591 bp in length
 * 106592 106691: gap of 100 bp
 * 106692 195080: contig of 88389 bp in length.
FEATURES Location/Qualifiers
 source
 1..195080
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /chromosome="25"
 /clone="CH211-218P2"
 /clone_lib="CHORI-211"
 1..106591
 /note="assembly_fragment:02685"
 106692..195080
 /note="assembly_fragment:04350"
ORIGIN
 Query Match 89.5%; Score 17; DB 12; Length 195080;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 TTTACTTCATAGTCTTT 18
 |||||
 Db 88539 TTTACTTCATAGTCTTT 88555
 RESULT 99
 AC111133

LOCUS AC111133 231443 bp DNA linear HTG 26-MAR-2004
DEFINITION Mus musculus chromosome 6 clone RP23-346C15 map 6, *** SEQUENCING
IN PROGRESS ***, 7 unordered pieces.
ACCESSION AC111133
VERSION AC111133.3 GI:45752893
KEYWORDS HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 231443)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 6, clone RP23-346C15
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 231443)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,
 MacDonald, P., Majot, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
 Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
 Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 231443)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 McLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R.,
 Seaman, S., Sever, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.
Direct Submission
 Submitted (26-MAR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 26, 2004 this sequence version replaced gi:30018044.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>
 Contact: sequence_submissions@road.mit.edu
 ----- Project Information
 Center project name: L22671
 Center clone name: 346_C15

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 107127: contig of 107127 bp in length
 * 107128: gap of 100 bp
 * 107228 120272: contig of 13045 bp in length
 * 120273 120372: gap of 100 bp
 * 120373 145416: contig of 25044 bp in length
 * 145417 145516: gap of 100 bp
 * 145517 166021: contig of 20505 bp in length
 * 166022 166121: gap of 100 bp
 * 166122 203366: contig of 37245 bp in length
 * 203367 203466: gap of 100 bp
 * 203467 208278: contig of 4812 bp in length
 * 208279 208378: gap of 100 bp
 * 208379 231443: contig of 23065 bp in length.

FEATURES

source
 1..231443
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="6"
 /map="6"
 /clone="RP23-346C15"
 /clone_lib="RPCI-23 Female Mouse BAC"
 107128..107227
 /estimated_length=100
 120273..120372
 /estimated_length=100
 145417..145516
 /estimated_length=100
 166022..166121
 /estimated_length=100
 203367..203466
 /estimated_length=100
 208279..208378
 /estimated_length=100

ORIGIN

Query Match 89.5%; Score 17; DB 12; Length 231443;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCTTT 18
 |||||

Db 195337 TTTACTTCATAGTCTTT 195353

RESULT 100

AC158667
 LOCUS AC158667 234055 bp DNA linear ROD 28-MAY-2005
 DEFINITION Mus musculus 6 BAC RP23-83P5 (Roswell Park Cancer Institute
 (C57Bl/6J Female) Mouse BAC Library) complete sequence.

ACCESSION AC158667

VERSION AC158667.8 GI:66792971

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 234055)

AUTHORS

Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsebrook, S., Archer, P.,
 Arredondo, H., Bandaranaike, D., Bangura, L., Beltzan, B., Beltran, R.,
 Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
 Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
 Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
 Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
 Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De
 Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
 Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
 Druada, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
 Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
 Fowler, G., Fu, O., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
 Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
 Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haebler, K.,
 Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
 Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
 Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,
 Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
 Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
 Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, P.,
 Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B.,
 Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
 London, P., Lopez, J., Lorensuhewa, L., Lozado, R., Luk, T., Madu, R.,
 Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
 McLelland, H., McPherson, J., Mercadado, C., Metzker, M.,
 Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
 Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastough, E.,
 Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okor, C., Odeh, E.,
 Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
 Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
 Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quirroz, J.,
 Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
 Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W.,
 Santibanez, J., Santos, R., Saverly, G., Scherer, S., Shen, H., Shen, Y.,
 Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
 Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
 Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villalana, D., Virk, D.,
 Volkov, A., Waldron, L., Walker, B., Wang, O., Wang, S., Warren, J.,
 Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
 Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
 Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 234055)
 Worley, K.C.
 Direct Submission
 Submitted (23-MAR-2005) Human Genome Sequencing Center, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 234055)
 Worley, K.C.
 Direct Submission
 Submitted (22-MAY-2005) Human Genome Sequencing Center, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 234055)
 Worley, K.C.
 Direct Submission
 Submitted (28-MAY-2005) Human Genome Sequencing Center, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 28, 2005 this sequence version replaced gi:66392614.
 Sequencing is completed to a minimum standard of double strand
 coverage with a minimum of 2 clones and 2 reads with no ambiguities
 or 2 chemistries with a minimum of 2 clones and 3 reads with no
 ambiguities. If the sequence quality does not meet this standard,
 it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
 Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
 Zhang.

Exon/Intron boundaries of identified genes were chosen if there
 were canonical splice junctions that maintained sequence continuity

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:16:40 ; Search time 141.233 Seconds
(without alignments)
937.971 Million cell updates/sec

Title: US-10-600-816-32

Perfect score: 19

Sequence: 1 ctttactcatgctttg 19

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : N Geneseq_8:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

14: Geneseqn2005s:*

15: Geneseqn2006s:*

Pred. No. is the number of results' predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	19	100.0	19	12	AD128529	Adi28529 Human GPC
2	19	100.0	497	11	ADT96418	Adt96418 Colon can
3	19	100.0	497	11	ADT96418	Adt96418 Colon can
4	19	100.0	552	11	ADT96292	Adt96292 Colon can
5	19	100.0	552	11	ADT96292	Adt96292 Colon can
6	19	100.0	603	3	AAZ90046	Aaz90046 Hydrophob
7	19	100.0	620	6	ABQ58527	Abq58527 Human col
8	19	100.0	634	6	ABQ58527	Abq58527 Human col
9	19	100.0	642	10	ABT22328	Abt22328 Breast ca
10	19	100.0	1074	12	ADO30035	Ado30035 Human GPC
11	19	100.0	1212	3	AAZ61776	Aaz61776 cDNA enco
12	19	100.0	1212	4	AAZ61776	Aaz61776 cDNA enco
13	19	100.0	1212	6	ABL34861	Ab134861 Human cell
14	19	100.0	1228	8	ACA04775	Aca04775 cDNA enco
15	19	100.0	1460	14	AEA00112	Aea00112 Human TAT
16	19	100.0	1460	14	AEA00112	Aea00112 Human TAT
17	19	100.0	1619	4	AAF58615	Aaf58615 Human REC
18	19	100.0	1718	3	AAZ90056	Aaz90056 Hydrophob

100.0	2302	6	ABT10173	Abt10173 Human bre
100.0	2302	8	ACC58386	Acc58386 Human GPC
100.0	2302	10	ADD93240	Add93240 RAIG1 cod
100.0	2302	10	ADL24773	Adl24773 Intestina
100.0	2316	10	ABT31923	Abt31923 Human bre
100.0	2446	4	AAH14688	Aah14688 Human cDN
100.0	2456	8	ABZ42832	Abz42832 Human G p
100.0	2456	10	ACC72695	Acc72695 Human can
100.0	2456	11	ADN39301	Adn39301 Cancer/an
100.0	2456	12	ADI28525	Adi28525 Human GPC
100.0	2456	12	ADI28459	Adi28459 Human GPC
100.0	2456	13	ADR48221	Adr48221 Human ret
100.0	2456	13	ACN39388	Acn39388 Tumour-as
100.0	2456	13	ADR43992	Adr43992 Human bre
100.0	2456	13	ADU06126	Adu06126 Novel bro
100.0	2456	14	AEA00088	Aea00088 Human TAT
100.0	2456	14	AEA00608	Aea00608 Human TAT
100.0	2456	14	AE47480	Aed47480 Retinoic
100.0	2593	6	ABQ54954	Abq54954 Human ova
100.0	3371	4	AAH72766	Aah72766 Human cer
100.0	4239	11	ACN89274	Acn89274 Breast ca
100.0	6730	13	ADX97494	Adx97494 Pancreat
91.6	1725	3	AAA96701	Aaa96701 Reporter
91.6	1788	10	ADF70573	Adf70573 Orphan re
91.6	188273	15	AEF80156	Aef80156 Cancer-as
86.3	168	14	AEF07706	Aef07706 Tobacco p
86.3	168	15	AEF98099	Aef98099 Tobacco c
86.3	354	3	AAA62561	Aaa62561 Human bre
86.3	354	6	ABT12676	Abt12676 Orestes s
86.3	354	10	ACD91970	Acd91970 Human col
86.3	407	3	AAK28686	Aak28686 Human sec
84.2	429	4	AAK74755	Aak74755 Human imm
84.2	491	4	AAK88468	Aak88468 Human dig
84.2	910	4	AAI97742	Aai97742 Human neu
84.2	981	4	AAI97741	Aai97741 Human neu
84.2	1596	4	AAK90456	Aak90456 Human dig
84.2	1596	4	AAK90457	Aak90457 Human dig
84.2	3447	15	AEF99265	Aef99265 Human EFH
84.2	3737	10	ADJ80212	Adj80212 Novel hum
84.2	3737	14	AEC20314	Aec20314 Human nuc
84.2	3739	3	AAK75928	Aak75928 Human ORF
84.2	3786	15	AEF22147	Aef22147 Human bas
84.2	4816	13	ACN39480	Acn39480 Tumour-as
84.2	127567	14	AEA61137	Aea61137 Human BCA
83.2	483	3	AAK36870	Aak36870 Arabidops
83.2	654	12	ADK16126	Adk16126 Nanoarcha
83.2	1131	12	ADN73034	Adn73034 Thale cre
83.2	1199	4	AAK25932	Aak25932 Human cDN
83.2	1258	10	ADB81647	Adb81647 Human ova
83.2	1259	8	ABX73273	Abx73273 Human nov
83.2	1366	8	ACC43801	Acc43801 Nucleotid
83.2	1394	3	AAK49655	Aak49655 Arabidops
83.2	1398	3	AAK33429	Aak33429 Arabidops
83.2	1401	4	AAI60916	Aai60916 Human pol
83.2	1599	9	ADB83252	Adb83252 Human cDN
83.2	1734	5	AAK64753	Aak64753 DNA enco
83.2	1758	11	ADN38895	Adn38895 Cancer/an
83.2	1868	8	ACC43802	Acc43802 Nucleotid
83.2	2000	12	ADJ41386	Adj41386 Plant cDN
83.2	2021	4	ABL60217	Ab160217 Human enc
83.2	2063	4	AAH16484	Aah16484 Human cDN
83.2	2094	4	ABL20545	Ab120545 Drosophil
83.2	2165	4	ABL05491	Ab105491 Drosophil
83.2	2175	6	ABQ72580	Abq72580 Human MDD
83.2	2208	14	ADZ26902	Adz26902 Adeno-ass
83.2	2208	14	ADZ26918	Adz26918 Adeno-ass
83.2	2208	14	ADZ26920	Adz26920 Adeno-ass
83.2	2208	14	ADZ26907	Adz26907 Adeno-ass
83.2	2208	14	ADZ26903	Adz26903 Adeno-ass
83.2	2208	14	ADZ26901	Adz26901 Adeno-ass
83.2	2208	14	ADZ26905	Adz26905 Adeno-ass
83.2	2208	14	ADZ26906	Adz26906 Adeno-ass
83.2	2208	14	ADZ26900	Adz26900 Adeno-ass

c 92	15.8	83.2	2208	14	ADZ26908	Adz26908 Adeno-ass	165	15.4	81.1	1825	5	ADL62252	Adl62252 Human ova
c 93	15.8	83.2	2208	14	ADZ26909	Adz26909 Adeno-ass	c 166	15.4	81.1	1941	8	ACA28183	Aca28183 Prokaryot
c 94	15.8	83.2	2208	14	ADZ26904	Adz26904 Adeno-ass	167	15.4	81.1	2000	11	ACL38594	Act38594 Rice stre
c 95	15.8	83.2	2208	14	ADZ26919	Adz26919 Adeno-ass	168	15.4	81.1	2028	14	AED20066	Aed20066 Human ova
c 96	15.8	83.2	2366	4	AAH72833	Aah72833 Human cer	169	15.4	81.1	2193	5	ABA21392	Ab21392 Human ner
c 97	15.8	83.2	2376	8	ACC43789	Acc43789 Nucleotid	170	15.4	81.1	2193	5	ABA21391	Ab21391 Human ner
c 98	15.8	83.2	2412	4	AAI59130	Aai59130 Human pol	171	15.4	81.1	2197	4	ABA09063	Ab209063 Human sec
c 99	15.8	83.2	2412	5	ADQ99353	Adq99353 DNA encod	c 172	15.4	81.1	2397	10	ABE262505	Abz262505 Human cdn
c 100	15.8	83.2	2412	9	ADB49113	Adb49113 Novel hum	173	15.4	81.1	2463	6	ABZ57269	Abz57269 Human pte
c 101	15.8	83.2	2438	5	ABV24578	Abv24578 Human pro	174	15.4	81.1	3299	3	AAZ36225	Aaz36225 A 3.3 kb
c 102	15.8	83.2	2438	5	ABV29178	Abv29178 Human pro	175	15.4	81.1	3442	12	ADQ74645	Adq74645 DNA of Ca
c 103	15.8	83.2	2441	4	ABL17716	Ab117716 Drosophil	176	15.4	81.1	3522	5	AAS86462	Aas86462 DNA encod
c 104	15.8	83.2	2459	5	AAS03040	Aas03040 Human dia	177	15.4	81.1	3638	2	AAX87602	Aax87602 Human sod
c 105	15.8	83.2	2488	5	AAS86426	Aas86426 DNA encod	178	15.4	81.1	3669	12	ADF29134	Adf29134 Human sec
c 106	15.8	83.2	4590	12	ADQ63855	Adq63855 Novel hum	179	15.4	81.1	3701	4	AAF30104	Aaf30104 Human sod
c 107	15.8	83.2	10195	4	AAS26626	Aas26626 Human gen	180	15.4	81.1	4013	11	ADM10936	Adm10936 Human O64
c 108	15.8	83.2	10195	8	ABX73975	Abx73975 Human nov	181	15.4	81.1	4013	12	ADJ11266	Adj11266 Human ova
c 109	15.8	83.2	12462	4	ABL05490	Ab105490 Drosophil	182	15.4	81.1	4013	12	ADM43527	Adm43527 Human ova
c 110	15.8	83.2	12481	4	ABL20544	Ab120544 Drosophil	c 183	15.4	81.1	4259	3	AAA74870	Aaa74870 Murine ch
c 111	15.8	83.2	27383	12	ADQ59497	Adq59497 Human can	c 184	15.4	81.1	4385	4	ABL07750	Ab107750 Drosophil
c 112	15.8	83.2	27898	14	ADZ13868	Adz13868 Murine ca	185	15.4	81.1	4385	3	AAA35009	Aaa35009 Human ade
c 113	15.8	83.2	83300	15	AEF07227	Aef07227 HRMT1L3 l	186	15.4	81.1	5262	14	AED29796	Aed29796 Human SCN
c 114	15.8	83.2	87394	12	ADT55151	Adt55151 Nucleotid	187	15.4	81.1	5475	14	AED29798	Aed29798 Human SCN
c 115	15.8	83.2	110000	12	ADK16049	Adk16049 Nanoarcha	188	15.4	81.1	5587	13	ADR06844	Adr06844 Full leng
c 116	15.8	83.2	127678	13	ABD32815	Abd32815 Mouse can	189	15.4	81.1	5728	6	AAI42750	Aai42750 Human sod
c 117	15.8	83.2	138251	13	ADQ80324	Adq80324 Human PAC	190	15.4	81.1	5860	4	AAF30101	Aaf30101 Human sod
c 118	15.8	83.2	169739	6	ABQ88186	Abq88186 Human ost	191	15.4	81.1	5860	10	ADD32193	Add32193 Human Na
c 119	15.8	83.2	209612	12	ADQ59395	Adq59395 Human can	192	15.4	81.1	6099	14	ADV50992	Adv50992 Human car
c 120	15.8	83.2	209612	14	ADZ13662	Adz13662 Murine ca	193	15.4	81.1	6099	14	ADV51070	Adv51070 Human car
c 121	15.8	83.2	235060	11	ACN43912	Acn43912 Mouse gen	194	15.4	81.1	6528	6	AAI42749	Aai42749 Human sod
c 122	15.8	83.2	265118	5	AAH41227	Aah41227 Pyrococcu	195	15.4	81.1	6528	14	ADL25905	Adl25905 Novel cel
c 123	15.8	83.2	337344	13	ABD32715	Abd32715 Human can	196	15.4	81.1	6870	14	ADV50994	Adv50994 Human fib
c 124	15.6	82.1	23	12	ADI28532	Adi28532 Human GPC	197	15.4	81.1	6870	14	ADV50996	Adv50996 Human fib
c 125	15.6	82.1	42	12	ADI28530	Adi28530 Human GPC	198	15.4	81.1	6870	14	ADV51074	Adv51074 Human fib
c 126	15.4	81.1	254	10	ACD95928	AcD95928 Human col	199	15.4	81.1	6870	14	ADV51072	Adv51072 Human fib
c 127	15.4	81.1	273	14	ADV51034	Adv51034 Human car	200	15.4	81.1	6988	9	ACD06169	AcD06169 Human cDN
c 128	15.4	81.1	273	14	ADV51112	Adv51112 Human car	201	15.4	81.1	7020	13	ADR90518	Adr90518 Human cDN
c 129	15.4	81.1	339	9	ACH31488	Ach31488 Human bon	202	15.4	81.1	7049	12	ADR95947	Adn95947 Human NOV
c 130	15.4	81.1	367	4	AAI11943	Aai11943 Human bre	203	15.4	81.1	7140	14	ADV51000	Adv51000 Human fib
c 131	15.4	81.1	368	4	AAI20833	Aai20833 Human bre	204	15.4	81.1	7140	14	ADV51078	Adv51078 Human fib
c 132	15.4	81.1	426	11	ACN82135	Acn82135 Breast ca	205	15.4	81.1	7143	14	ADV51002	Adv51002 Human fib
c 133	15.4	81.1	451	6	ABT06979	Abt06979 Human ova	206	15.4	81.1	7143	14	ADV51080	Adv51080 Human fib
c 134	15.4	81.1	451	8	ABX72857	Abx72857 Human ova	207	15.4	81.1	7161	14	ADM63878	Adm63878 Human fib
c 135	15.4	81.1	455	4	AAI12184	Aai12184 Probe #21	208	15.4	81.1	7242	13	ADR66637	Adr66637 Human pro
c 136	15.4	81.1	455	4	ABA53889	Ab253889 Human foe	209	15.4	81.1	7242	13	ADR65953	Adr65953 Human pro
c 137	15.4	81.1	455	4	AAI33530	Aai33530 Probe #22	210	15.4	81.1	7361	9	ACD06170	AcD06170 Human cDN
c 138	15.4	81.1	455	4	ABA43438	Ab243438 Human bre	211	15.4	81.1	7361	12	ADR95949	Adn95949 Human NOV
c 139	15.4	81.1	455	4	ABA23636	Ab223636 Probe #21	212	15.4	81.1	7388	15	AEF38754	Aef38754 Human fib
c 140	15.4	81.1	455	4	AAK27602	Aak27602 Human bon	213	15.4	81.1	7388	15	AEF38715	Aef38715 Human fib
c 141	15.4	81.1	455	4	AAK02154	Aak02154 Human bra	214	15.4	81.1	7388	15	AEF35211	Aef35211 Human fib
c 142	15.4	81.1	455	4	AAI27173	Aai27173 Human liv	215	15.4	81.1	7388	15	AEF64032	Aef64032 Human fib
c 143	15.4	81.1	455	5	AAI02097	Aai02097 Probe #20	216	15.4	81.1	7388	15	AEF63949	Aef63949 Human fib
c 144	15.4	81.1	455	6	ABS02070	Ab202070 Human gen	217	15.4	81.1	7413	14	ADV50998	Adv50998 Human fib
c 145	15.4	81.1	456	9	ACH40353	Ach40353 Human foe	218	15.4	81.1	7413	14	ADV51076	Adv51076 Human car
c 146	15.4	81.1	482	9	ACH41862	Ach41862 Human foe	219	15.4	81.1	7550	8	ACC00412	Acc00412 Human cel
c 147	15.4	81.1	522	8	ACF73000	Acf73000 Staphyloc	220	15.4	81.1	7677	13	ACN37820	Acn37820 Tumour-as
c 148	15.4	81.1	523	14	AEI12931	Aei12931 Hamster c	221	15.4	81.1	7679	10	ADZ31322	Adz31322 Testoster
c 149	15.4	81.1	523	14	AEI16573	Aei16573 Hamster S	222	15.4	81.1	7679	14	ADZ75397	Adz75397 Human fib
c 150	15.4	81.1	591	4	AAI16155	Aai16155 Human bre	223	15.4	81.1	7680	2	AAI17551	Aai17551 Human fib
c 151	15.4	81.1	605	10	ABT21654	Abt21654 Breast ca	224	15.4	81.1	7680	3	AAI21131	Aai21131 Human low
c 152	15.4	81.1	611	11	ACN92521	Acn92521 Breast ca	225	15.4	81.1	7680	5	ABA82689	Ab282689 Fibronec
c 153	15.4	81.1	773	4	AAI24998	Aai24998 Human bre	226	15.4	81.1	7680	6	ABL67540	Ab167540 Thyroid c
c 154	15.4	81.1	845	11	ACN86103	Acn86103 Breast ca	227	15.4	81.1	7680	6	ABT11082	Abt11082 Human bre
c 155	15.4	81.1	856	2	AAI60242	Aai60242 cDNA enco	228	15.4	81.1	7680	8	ACF03878	Acf03878 Human fib
c 156	15.4	81.1	873	11	ADM10903	Adm10903 Human ova	229	15.4	81.1	7680	8	ABX10391	Abx10391 DNA encod
c 157	15.4	81.1	873	12	ADJ11233	Adj11233 Human ova	230	15.4	81.1	7680	8	ACC46009	Acc46009 Human fib
c 158	15.4	81.1	873	12	ADMA3494	Adm3494 Human ova	231	15.4	81.1	7680	8	ACF12859	Acf12859 Human cer
c 159	15.4	81.1	927	8	ACA24600	Aca24600 Prokaryot	232	15.4	81.1	7680	8	ACA64817	Ac264817 Human fib
c 160	15.4	81.1	1070	13	ADX49239	Adx49239 Plant ful	233	15.4	81.1	7680	9	ADB70377	Adb70377 Fibronec
c 161	15.4	81.1	1312	5	ABA14166	Ab214166 Human ner	234	15.4	81.1	7680	10	ADB98703	Adb98703 Human fib
c 162	15.4	81.1	1518	10	ACF72171	Acf72171 Photorbab	235	15.4	81.1	7680	10	ADD18771	Add18771 Human dis
c 163	15.4	81.1	1562	11	ACN90569	Acn90569 Breast ca	236	15.4	81.1	7680	10	ADZ82499	Adz82499 Human DNA
c 164	15.4	81.1	1806	13	ADS57710	Ads57710 Bacterial	237	15.4	81.1	7680	10	ADZ82499	Adz82499 Human DNA

238	15.4	81.1	7680	10	ABZ96825	Human nuc
239	15.4	81.1	7680	11	AD1311963	Human cDN
240	15.4	81.1	7680	11	ABD20674	Human pul
241	15.4	81.1	7680	12	ADJ37156	Human mal
242	15.4	81.1	7680	12	ADP33486	Renal cel
243	15.4	81.1	7680	13	ADR99073	Fibronect
244	15.4	81.1	7680	13	ADS84030	Human lym
245	15.4	81.1	7680	14	ADV70128	Tumor-ass
246	15.4	81.1	7680	14	ACL53987	Human col
247	15.4	81.1	7680	14	ADZ36441	Human fib
248	15.4	81.1	7680	14	ABE11753	Human fib
249	15.4	81.1	7705	1	AAH70596	Sequence
250	15.4	81.1	7705	2	AAQ05502	Sequence
251	15.4	81.1	7705	2	AAQ15214	Human fib
252	15.4	81.1	7705	12	ADG89561	Human fib
253	15.4	81.1	7710	12	ADQ28933	Mouse nov
254	15.4	81.1	7795	10	ADJ36196	Mouse nov
255	15.4	81.1	7803	2	ADQ70009	Fibrinoge
256	15.4	81.1	7848	13	ADQ38583	Human SNP
257	15.4	81.1	7867	4	AAI57803	Human pol
258	15.4	81.1	7912	14	ADZ26740	Human fib
259	15.4	81.1	7912	15	AEF38777	Oncofetal
260	15.4	81.1	7912	15	AEF38738	Oncofetal
261	15.4	81.1	7912	15	AEF35234	Human onc
262	15.4	81.1	7912	15	AEF64055	Human onc
263	15.4	81.1	7912	15	AEF63972	Human onc
264	15.4	81.1	7935	13	ADQ38577	Human SNP
265	15.4	81.1	7959	13	ADQ38586	Human SNP
266	15.4	81.1	8013	13	ADQ38579	Human SNP
267	15.4	81.1	8027	11	ADP64998	Human fib
268	15.4	81.1	8027	12	ADG89565	Human fib
269	15.4	81.1	8027	12	ADQ29601	Human col
270	15.4	81.1	8027	13	ADR67200	Human bla
271	15.4	81.1	8027	13	ADV70129	Tumor-ass
272	15.4	81.1	8027	14	AEA04397	Human cDN
273	15.4	81.1	8027	14	AEF17605	Human fib
274	15.4	81.1	8027	15	AEF69913	Colorecta
275	15.4	81.1	8039	3	AAAC89889	Human FN
276	15.4	81.1	8044	6	AAAC89886	Human DNA
277	15.4	81.1	8044	12	ADI61767	Human cDN
278	15.4	81.1	8044	14	AEA43932	Human cDN
279	15.4	81.1	8062	4	AAI57802	Human pol
280	15.4	81.1	8066	4	AAI59591	Human pol
281	15.4	81.1	8066	4	AAI59589	Human pol
282	15.4	81.1	8066	4	AAI59588	Human pol
283	15.4	81.1	8066	4	AAI59590	Human pol
284	15.4	81.1	8137	4	AAI57805	Human pol
285	15.4	81.1	8155	13	ADQ38584	Human SNP
286	15.4	81.1	8216	5	AAAC86466	DNA encod
287	15.4	81.1	8226	13	ADQ38574	Human SNP
288	15.4	81.1	8230	4	AAI57804	Human pol
289	15.4	81.1	8232	14	AEF73865	Human pla
290	15.4	81.1	8272	14	ADZ26744	Human fib
291	15.4	81.1	8272	15	AEF38775	Oncofetal
292	15.4	81.1	8272	15	AEF38736	Oncofetal
293	15.4	81.1	8272	15	AEF35232	Human onc
294	15.4	81.1	8272	15	AEF64053	Human onc
295	15.4	81.1	8272	15	AEF63970	Human onc
296	15.4	81.1	8278	13	ADQ38587	Human SNP
297	15.4	81.1	8332	13	ADQ38576	Human SNP
298	15.4	81.1	8371	13	ADQ38580	Human SNP
299	15.4	81.1	8374	14	ADZ26748	Human fib
300	15.4	81.1	8374	15	AEF38773	Oncofetal
301	15.4	81.1	8374	15	AEF38734	Oncofetal
302	15.4	81.1	8374	15	AEF35230	Human onc
303	15.4	81.1	8374	15	AEF64051	Human onc
304	15.4	81.1	8374	15	AEF63968	Human onc
305	15.4	81.1	8449	14	ADZ26564	Human fib
306	15.4	81.1	8449	15	AEF38771	Oncofetal
307	15.4	81.1	8449	15	AEF38732	Oncofetal
308	15.4	81.1	8449	15	AEF35228	Human onc
309	15.4	81.1	8449	15	AEF64049	Human onc
310	15.4	81.1	8449	15	AEF63966	Human onc
311	15.4	81.1	8647	14	ADZ26742	Human fib
312	15.4	81.1	8647	15	AEF38769	Oncofetal
313	15.4	81.1	8647	15	AEF38730	Oncofetal
314	15.4	81.1	8647	15	AEF35226	Human onc
315	15.4	81.1	8647	15	AEF64047	Human onc
316	15.4	81.1	8647	15	AEF63964	Human onc
317	15.4	81.1	8815	14	ADW81010	Fibronect
318	15.4	81.1	8815	14	ADX69342	DNA encod
319	15.4	81.1	8815	14	ADZ26746	Human fib
320	15.4	81.1	8815	15	AEF38767	Oncofetal
321	15.4	81.1	8815	15	AEF38728	Oncofetal
322	15.4	81.1	8815	15	AEF35224	Human onc
323	15.4	81.1	8815	15	AEF64045	Human onc
324	15.4	81.1	8815	15	AEF63962	Human onc
325	15.4	81.1	14698	4	AAAC27769	DNA encod
326	15.4	81.1	14698	8	ABZ74520	Secreted
327	15.4	81.1	14698	10	ADB94572	Novel hum
328	15.4	81.1	14698	10	ABZ68054	Human sec
329	15.4	81.1	14740	3	AAA35010	Human ade
330	15.4	81.1	14740	3	AAAF21132	Human low
331	15.4	81.1	14740	10	ABZ96826	Human nuc
332	15.4	81.1	14740	11	ABD20675	Human pul
333	15.4	81.1	17815	14	AED29845	pCMR17 pl
334	15.4	81.1	18031	14	AED29848	pCMR18 pl
335	15.4	81.1	24477	14	AEF39174	Continnuation (10 o
336	15.4	81.1	31749	4	AAK72959	Human imm
337	15.4	81.1	47583	13	ADT77145	Type II d
338	15.4	81.1	52938	12	ADQ59443	Continuation (4 of
339	15.4	81.1	52938	14	ADZ13754	Continuation (4 of
340	15.4	81.1	78925	3	AAAC89888	Human FN
341	15.4	81.1	98606	11	ACN43868	Mouse gen
342	15.4	81.1	110000	2	AAV21209	Continuation (13 o
343	15.4	81.1	110000	2	AAV21209	Continuation (14 o
344	15.4	81.1	110000	2	AAV22305	Continuation (5 of
345	15.4	81.1	110000	6	ABA90521	Continuation (21 o
346	15.4	81.1	110000	10	ACF67367	Continuation (55 o
347	15.4	81.1	110000	10	ACF65387	Continuation (6 of
348	15.4	81.1	110000	14	ADZ42285	Continuation (2 of
349	15.4	81.1	110000	14	AEBA2737	Continuation (11 o
350	15.4	81.1	143947	15	AEF38790	Human fib
351	15.4	81.1	143947	15	AEF38751	Human fib
352	15.4	81.1	143947	15	AEF35247	Human fib
353	15.4	81.1	143947	15	AEF64068	Human fib
354	15.4	81.1	143947	15	AEF63985	Human fib
355	15.4	81.1	150397	14	AEEO4865	Cancer-as
356	15.4	81.1	174318	14	AEF92655	Human Ras
357	15.4	81.1	199719	14	AEEO5001	Cancer-as
358	15.4	81.1	199994	14	AEA07496	Cancer-as
359	15.4	81.1	200000	12	ADO47190	DNA seque
360	15	78.9	40	12	ADI28527	Human GPC
361	15	78.9	158	4	ABA70485	Human foe
362	15	78.9	158	4	AAI50643	Probe #19
363	15	78.9	158	4	AAK44663	Human bon
364	15	78.9	158	4	AAK18730	Human bra
365	15	78.9	158	4	ABSA4323	Human liv
366	15	78.9	158	6	ABS18900	Human gen
367	15	78.9	476	4	ABA57875	Human foe
368	15	78.9	476	4	AAI37467	Probe #61
369	15	78.9	476	4	AAK31588	Human bon
370	15	78.9	476	4	AAK05948	Human bra
371	15	78.9	476	4	ABSA1272	Human liv
372	15	78.9	476	6	ABS06342	Human gen
373	15	78.9	487	10	ADE81977	Arabidops
374	15	78.9	599	12	ACH66907	Human gen
375	15	78.9	1155	6	ABZ13845	Arabidops
376	15	78.9	1155	8	ADA68289	Arabidops
377	15	78.9	1183	11	ACN89691	Breast ca
378	15	78.9	1439	3	AAC47707	Arabidops
379	15	78.9	1442	3	AAC41748	Arabidops
380	15	78.9	2262	3	AAZ59466	Human MPR
381	15	78.9	2415	11	ADL95396	Angiotens
382	15	78.9	2418	3	AAZ59465	Human MPR
383	15	78.9	2911	4	AAAS14890	Human cDN

c 384	15	78.9	2920	4	AAS14880	Aas14880 Human cDN	c 457	15	78.9	3732	9	ADA86651	Ada86651 Novel hum
c 385	15	78.9	2920	8	ADA03339	Ada03339 Human ang	c 458	15	78.9	3732	9	ADA87754	Ada87754 Novel hum
c 386	15	78.9	2920	8	ACC79021	Acc79021 Human ACE	c 459	15	78.9	3732	9	ADA46142	Ada46142 Novel hum
c 387	15	78.9	3064	13	ACN43792	Acn43792 Human dia	c 460	15	78.9	3732	9	ADB28172	Adb28172 cDNA enco
c 388	15	78.9	3072	13	ACN43791	Acn43791 Human dia	c 461	15	78.9	3732	9	ADB28724	Adb28724 cDNA enco
c 389	15	78.9	3324	9	ACH03967	Ach03967 Human cDN	c 462	15	78.9	3732	9	ADA76676	Ada76676 Human PRO
c 390	15	78.9	3325	10	ADC38727	Adc38727 Human cDN	c 463	15	78.9	3732	9	ADA88306	Ada88306 Novel hum
c 391	15	78.9	3334	4	AAC84366	Aac84366 Human Zac	c 464	15	78.9	3732	9	ADA97311	Ada97311 Human PRO
c 392	15	78.9	3334	10	ABX93333	Abx93333 cDNA enco	c 465	15	78.9	3732	9	ADB27068	Adb27068 cDNA enco
c 393	15	78.9	3396	3	AAAL2764	Aaal2764 cDNA enco	c 466	15	78.9	3732	9	ADB22001	Adb22001 Novel hum
c 394	15	78.9	3396	6	AAD02758	Aad02758 Human ang	c 467	15	78.9	3732	9	ADA66692	Ada66692 Human PRO
c 395	15	78.9	3396	6	AAD32586	Aad32586 Human ACE	c 468	15	78.9	3732	9	ADB22553	Adb22553 Human PRO
c 396	15	78.9	3396	6	ABK87623	Abk87623 cDNA enco	c 469	15	78.9	3732	9	ADB23326	Adb23326 Human PRO
c 397	15	78.9	3396	8	ADA03343	Ada03343 Human ang	c 470	15	78.9	3732	9	ADA92048	Ada92048 Human PRO
c 398	15	78.9	3396	8	ACC79022	Acc79022 Human ACE	c 471	15	78.9	3732	9	ADB15111	Adb15111 Human PRO
c 399	15	78.9	3396	11	ADL95334	Adl95334 Human ang	c 472	15	78.9	3732	9	ADB38363	Adb38363 Novel hum
c 400	15	78.9	3404	12	ADH51356	Adh51356 Human ang	c 473	15	78.9	3732	9	ADB37811	Adb37811 Novel hum
c 401	15	78.9	3405	6	ABS60632	Abs60632 Human cDN	c 474	15	78.9	3732	9	ADB66283	Adb66283 Novel hum
c 402	15	78.9	3405	6	ABS60371	Abs60371 Human cDN	c 475	15	78.9	3732	10	ADB89363	Adb89363 Human PRO
c 403	15	78.9	3405	14	AEC76234	Aec76234 Human ACE	c 476	15	78.9	3732	10	ADB90095	Adb90095 Human PRO
c 404	15	78.9	3405	14	AED14037	Aed14037 Human cum	c 477	15	78.9	3732	10	ADB39196	Adb39196 Novel hum
c 405	15	78.9	3474	5	AAS42515	Aas42515 Human cDN	c 478	15	78.9	3732	10	ADB46819	Adb46819 Novel hum
c 406	15	78.9	3732	4	AAS21279	Aas21279 Human cDN	c 479	15	78.9	3732	10	ADB86426	Adb86426 Human PRO
c 407	15	78.9	3732	8	ACA03638	Aca03638 cDNA enco	c 480	15	78.9	3732	10	ADB77031	Adb77031 Novel hum
c 408	15	78.9	3732	8	ABX89176	Abx89176 DNA encod	c 481	15	78.9	3732	10	ADB34188	Adb34188 Human PRO
c 409	15	78.9	3732	8	ACD41830	Acd41830 Human sec	c 482	15	78.9	3732	10	ADB35292	Adb35292 Human PRO
c 410	15	78.9	3732	8	ACA04059	Aca04059 Human cDN	c 483	15	78.9	3732	10	ADB33636	Adb33636 Human PRO
c 411	15	78.9	3732	9	ADA45590	Ada45590 Novel hum	c 484	15	78.9	3732	10	ADB34740	Adb34740 Human PRO
c 412	15	78.9	3732	9	ADA76021	Ada76021 Human PRO	c 485	15	78.9	3732	10	ADB35844	Adb35844 Human PRO
c 413	15	78.9	3732	9	ADA18671	Ada18671 Human PRO	c 486	15	78.9	3732	10	ADB46239	Adb46239 Novel hum
c 414	15	78.9	3732	9	ADA61294	Ada61294 Homo sapi	c 487	15	78.9	3732	10	ADC50112	Adc50112 Novel hum
c 415	15	78.9	3732	9	ADB19079	Adb19079 Novel hum	c 488	15	78.9	3732	10	ADC71659	Adc71659 Novel hum
c 416	15	78.9	3732	9	ADB27620	Adb27620 cDNA enco	c 489	15	78.9	3732	10	ADC59638	Adc59638 Novel hum
c 417	15	78.9	3732	9	ADA86099	Ada86099 Novel hum	c 490	15	78.9	3732	10	ADC52645	Adc52645 Novel hum
c 418	15	78.9	3732	9	ADB15663	Adb15663 Human PRO	c 491	15	78.9	3732	10	ADC56999	Adc56999 Novel hum
c 419	15	78.9	3732	9	ADA47449	Ada47449 Human PRO	c 492	15	78.9	3732	10	ADC60190	Adc60190 Novel hum
c 420	15	78.9	3732	9	ADA67244	Ada67244 Human PRO	c 493	15	78.9	3732	10	ADC50665	Adc50665 Novel hum
c 421	15	78.9	3732	9	ADB30251	Adb30251 cDNA enco	c 494	15	78.9	3732	10	ADC65192	Adc65192 Human PRO
c 422	15	78.9	3732	9	ADA85547	Ada85547 Novel hum	c 495	15	78.9	3732	10	ADC54290	Adc54290 Novel hum
c 423	15	78.9	3732	9	ADA96759	Ada96759 Human PRO	c 496	15	78.9	3732	10	ADC53251	Adc53251 Novel hum
c 424	15	78.9	3732	9	ADA79063	Ada79063 Human PRO	c 497	15	78.9	3732	10	ADC58774	Adc58774 Novel hum
c 425	15	78.9	3732	9	ADA87202	Ada87202 Novel hum	c 498	15	78.9	3732	10	ADC55652	Adc55652 Novel hum
c 426	15	78.9	3732	9	ADB16404	Adb16404 Human PRO	c 499	15	78.9	3732	10	ADC58222	Adc58222 Novel hum
c 427	15	78.9	3732	9	ADA91496	Ada91496 Novel hum	c 500	15	78.9	3732	10	ADC02896	Adc02896 Novel hum
c 428	15	78.9	3732	9	ADA14559	Ada14559 Human PRO	c 501	15	78.9	3732	10	ADC89888	Adc89888 Novel hum
c 429	15	78.9	3732	9	ADB18520	Adb18520 Novel hum	c 502	15	78.9	3732	10	ADC69307	Adc69307 cDNA enco
c 430	15	78.9	3732	9	ADA93735	Ada93735 Human PRO	c 503	15	78.9	3732	10	ADC48196	Adc48196 Human PRO
c 431	15	78.9	3732	9	ADB19631	Adb19631 Novel hum	c 504	15	78.9	3732	10	ADD009725	Add009725 Human PRO
c 432	15	78.9	3732	9	ADB12943	Adb12943 Human PRO	c 505	15	78.9	3732	10	ADD04300	Add04300 Novel hum
c 433	15	78.9	3732	9	ACD98459	Acd98459 Novel hum	c 506	15	78.9	3732	10	ADC80256	Adc80256 Novel hum
c 434	15	78.9	3732	9	ADA74197	Ada74197 Human PRO	c 507	15	78.9	3732	10	ADD10763	Add10763 Human PRO
c 435	15	78.9	3732	9	ADB24430	Adb24430 Human PRO	c 508	15	78.9	3732	10	ADC47644	Adc47644 Human PRO
c 436	15	78.9	3732	9	ADA81954	Ada81954 Human PRO	c 509	15	78.9	3732	10	ADC79704	Adc79704 Novel hum
c 437	15	78.9	3732	9	ADA74917	Ada74917 Human PRO	c 510	15	78.9	3732	10	ADD09173	Add09173 Human PRO
c 438	15	78.9	3732	9	ADA84995	Ada84995 Novel hum	c 511	15	78.9	3732	10	ADD40886	Add40886 Novel hum
c 439	15	78.9	3732	9	ADA84443	Ada84443 Novel hum	c 512	15	78.9	3732	10	ADD52025	Add52025 cDNA enco
c 440	15	78.9	3732	9	ADB29699	Adb29699 cDNA enco	c 513	15	78.9	3732	10	ADD52765	Add52765 cDNA enco
c 441	15	78.9	3732	9	ADA80227	Ada80227 Human PRO	c 514	15	78.9	3732	10	ADD53317	Add53317 Novel hum
c 442	15	78.9	3732	9	ADA75469	Ada75469 Human PRO	c 515	15	78.9	3732	10	ADD51473	Add51473 cDNA enco
c 443	15	78.9	3732	9	ADA46694	Ada46694 Human PRO	c 516	15	78.9	3732	10	ADD02272	Add02272 Human PRO
c 444	15	78.9	3732	9	ADB24990	Adb24990 Human PRO	c 517	15	78.9	3732	10	ADD01706	Add01706 Human PRO
c 445	15	78.9	3732	9	ADA93166	Ada93166 Human PRO	c 518	15	78.9	3732	10	ADD53888	Add53888 Novel hum
c 446	15	78.9	3732	9	ADB26516	Adb26516 cDNA enco	c 519	15	78.9	3732	10	ADD92205	Add92205 Human PRO
c 447	15	78.9	3732	9	ADB30803	Adb30803 cDNA enco	c 520	15	78.9	3732	10	ADD91101	Add91101 Human PRO
c 448	15	78.9	3732	9	ADA60731	Ada60731 Homo sapi	c 521	15	78.9	3732	10	ADE03715	Ade03715 Human PRO
c 449	15	78.9	3732	9	ADB23878	Adb23878 Human PRO	c 522	15	78.9	3732	10	ADE32012	Ade32012 Novel hum
c 450	15	78.9	3732	9	ADA96207	Ada96207 Human PRO	c 523	15	78.9	3732	10	ADE21944	Ade21944 cDNA enco
c 451	15	78.9	3732	9	ADA80779	Ada80779 Human PRO	c 524	15	78.9	3732	10	ADD79168	Add79168 cDNA enco
c 452	15	78.9	3732	9	ADA95655	Ada95655 Human PRO	c 525	15	78.9	3732	10	ADE41704	Ade41704 Human PRO
c 453	15	78.9	3732	9	ADB25964	Adb25964 cDNA enco	c 526	15	78.9	3732	10	ADE17521	Ade17521 Human PRO
c 454	15	78.9	3732	9	ADB21449	Adb21449 Novel hum	c 527	15	78.9	3732	10	ADD91653	Add91653 Human PRO
c 455	15	78.9	3732	9	ADA77228	Ada77228 Human PRO	c 528	15	78.9	3732	10	ADB33116	Adb33116 Novel hum
c 456	15	78.9	3732	9	ADB17968	Adb17968 cDNA enco	c 529	15	78.9	3732	10	ADB33668	Adb33668 Novel hum

c 530	15	78.9	3732	10	ADD79720	Adh79720 cDNA enco	c 603	15	78.9	3732	12	ADG06901	Novel hum
c 531	15	78.9	3732	10	ADD92757	Adg92757 Human PRO	c 604	15	78.9	3732	12	ADG07453	Novel hum
c 532	15	78.9	3732	10	ADe19177	Human PRO	c 605	15	78.9	3732	12	ADG54948	Novel hum
c 533	15	78.9	3732	10	ADe18625	Human PRO	c 606	15	78.9	3732	12	ADG60612	Novel hum
c 534	15	78.9	3732	10	ADe42821	Human PRO	c 607	15	78.9	3732	12	ADG61716	Novel hum
c 535	15	78.9	3732	10	ADd95610	Human PRO	c 608	15	78.9	3732	12	ADG81917	Human PRO
c 536	15	78.9	3732	10	ADd22496	cDNA enco	c 609	15	78.9	3732	12	ADG57156	Novel hum
c 537	15	78.9	3732	10	ADd78614	cDNA enco	c 610	15	78.9	3732	12	ADG56604	Novel hum
c 538	15	78.9	3732	10	ADe32564	Novel hum	c 611	15	78.9	3732	12	ADG55500	Novel hum
c 539	15	78.9	3732	10	ADd42256	Human PRO	c 612	15	78.9	3732	12	ADG58260	Novel hum
c 540	15	78.9	3732	10	ADd80272	cDNA enco	c 613	15	78.9	3732	12	ADG70626	Novel hum
c 541	15	78.9	3732	10	ADd89300	Human PRO	c 614	15	78.9	3732	12	ADG57708	Novel hum
c 542	15	78.9	3732	10	ADe40584	Human PRO	c 615	15	78.9	3732	12	ADG53292	Novel hum
c 543	15	78.9	3732	10	ADe04383	Human PRO	c 616	15	78.9	3732	12	ADG71178	Novel hum
c 544	15	78.9	3732	10	ADe92512	Human PRO	c 617	15	78.9	3732	12	ADG811365	Human PRO
c 545	15	78.9	3732	10	ADG21221	Novel hum	c 618	15	78.9	3732	12	ADH30327	Human PRO
c 546	15	78.9	3732	10	ADG22862	Novel hum	c 619	15	78.9	3732	12	ADH11694	Novel hum
c 547	15	78.9	3732	10	ADf97197	Human PRO	c 620	15	78.9	3732	12	ADG52116	Novel hum
c 548	15	78.9	3732	10	ADG80261	Human PRO	c 621	15	78.9	3732	12	ADG53844	Novel hum
c 549	15	78.9	3732	10	ADG79709	Human PRO	c 622	15	78.9	3732	12	ADG80813	Human PRO
c 550	15	78.9	3732	10	ADH55001	Novel hum	c 623	15	78.9	3732	12	ADG56052	Novel hum
c 551	15	78.9	3732	10	ADH55553	Novel hum	c 624	15	78.9	3732	12	ADH12318	Novel hum
c 552	15	78.9	3732	10	ADi64721	Novel hum	c 625	15	78.9	3732	12	ADG61164	Novel hum
c 553	15	78.9	3732	10	ADH81634	Novel hum	c 626	15	78.9	3732	12	ADH28251	Human PRO
c 554	15	78.9	3732	10	ADH81082	Novel hum	c 627	15	78.9	3732	12	ADG54396	Novel hum
c 555	15	78.9	3732	10	ACd23888	Novel hum	c 628	15	78.9	3732	12	ADG59436	Novel hum
c 556	15	78.9	3732	10	ACa67029	cDNA enco	c 629	15	78.9	3732	12	ADi180860	cDNA enco
c 557	15	78.9	3732	11	ADH82251	Novel hum	c 630	15	78.9	3732	12	ADG09603	Novel hum
c 558	15	78.9	3732	11	ADN15650	Novel hum	c 631	15	78.9	3732	12	ADi15074	Novel hum
c 559	15	78.9	3732	11	ADN16279	Novel hum	c 632	15	78.9	3732	12	ADG08951	Novel hum
c 560	15	78.9	3732	11	ADN15098	Novel hum	c 633	15	78.9	3732	12	ADi14406	Novel hum
c 561	15	78.9	3732	11	ADN14546	Novel hum	c 634	15	78.9	3732	12	ADi18001	Novel hum
c 562	15	78.9	3732	11	ADi63772	Novel hum	c 635	15	78.9	3732	12	ADJ63282	Novel hum
c 563	15	78.9	3732	11	ADi63220	Novel hum	c 636	15	78.9	3732	12	ADJ77177	Human PRO
c 564	15	78.9	3732	12	ADc80808	Novel hum	c 637	15	78.9	3732	12	ADJ65299	cDNA enco
c 565	15	78.9	3732	12	ADd76256	Human PRO	c 638	15	78.9	3732	12	ADM27435	cDNA enco
c 566	15	78.9	3732	12	ADd87620	Human PRO	c 639	15	78.9	3732	12	ADM42159	cDNA enco
c 567	15	78.9	3732	12	ADd86024	Human PRO	c 640	15	78.9	3732	12	ADM28021	cDNA enco
c 568	15	78.9	3732	12	ADe75472	Human PRO	c 641	15	78.9	3732	13	ADi95503	cDNA enco
c 569	15	78.9	3732	12	ADe23048	cDNA enco	c 642	15	78.9	3732	13	ADi96055	Novel hum
c 570	15	78.9	3732	12	ADe23600	cDNA enco	c 643	15	78.9	3732	13	ADi96055	Novel hum
c 571	15	78.9	3732	12	ADe24243	cDNA enco	c 644	15	78.9	3732	13	ADT02991	Human PRO
c 572	15	78.9	3732	12	ADd87068	Human PRO	c 645	15	78.9	3732	14	ADB203042	Human sec
c 573	15	78.9	3732	12	ADd88934	Human PRO	c 646	15	78.9	3732	14	AEb13788	Cancer ce
c 574	15	78.9	3732	12	ADe18073	Human PRO	c 647	15	78.9	3732	14	AEC10071	Human PRO
c 575	15	78.9	3732	12	ADd88382	Human PRO	c 648	15	78.9	3732	14	AEC76242	Human ACE
c 576	15	78.9	3732	12	ADe94402	cDNA enco	c 649	15	78.9	3732	14	ADe85986	Human PRO
c 577	15	78.9	3732	12	ADe90813	Human PRO	c 650	15	78.9	40782	14	AEC76241	Human X-c
c 578	15	78.9	3732	12	ADe94954	cDNA enco	c 651	15	78.9	76138	12	ADQ97334	Mouse can
c 579	15	78.9	3732	12	ADe93064	Human PRO	c 652	14.8	77.9	144	12	ADQ04602	Maize hom
c 580	15	78.9	3732	12	ADf34645	cDNA enco	c 653	14.8	77.9	156	14	AEb89652	Isolated
c 581	15	78.9	3732	12	ADe91960	Novel hum	c 654	14.8	77.9	175	13	ACF89480	Human SIR
c 582	15	78.9	3732	12	ADe90261	Human PRO	c 655	14.8	77.9	238	10	ADF81263	Leukaemia
c 583	15	78.9	3732	12	ADf98320	Human PRO	c 656	14.8	77.9	247	3	AAF15326	Trichoder
c 584	15	78.9	3732	12	ADG01987	Human PRO	c 657	14.8	77.9	247	13	ADU59367	Trichoder
c 585	15	78.9	3732	12	ADG21773	Novel hum	c 658	14.8	77.9	247	14	ADZ97370	Trichoder
c 586	15	78.9	3732	12	ADG19843	cDNA enco	c 659	14.8	77.9	249	8	ACA23185	Prokaryot
c 587	15	78.9	3732	12	ADf97749	Human PRO	c 660	14.8	77.9	266	2	AAT21096	Human gen
c 588	15	78.9	3732	12	ADG23966	Novel hum	c 661	14.8	77.9	295	3	AAA69633	Eucalyptu
c 589	15	78.9	3732	12	ADf98320	Human PRO	c 662	14.8	77.9	300	2	AAZ14568	Human gen
c 590	15	78.9	3732	12	ADG03151	Human PRO	c 663	14.8	77.9	306	10	ABX78380	Soybean s
c 591	15	78.9	3732	12	ADf98872	Human PRO	c 664	14.8	77.9	306	15	AEF94242	Stress re
c 592	15	78.9	3732	12	ADG16457	cDNA enco	c 665	14.8	77.9	326	2	AAQ80918	Spruce tr
c 593	15	78.9	3732	12	ADG04916	Human PRO	c 666	14.8	77.9	337	4	AAK78974	Human imm
c 594	15	78.9	3732	12	ADG19183	cDNA enco	c 667	14.8	77.9	337	4	AAK78975	Human imm
c 595	15	78.9	3732	12	ADG13020	cDNA enco	c 668	14.8	77.9	352	4	AAK60950	Human imm
c 596	15	78.9	3732	12	ADG08077	Novel hum	c 669	14.8	77.9	363	3	AAC31404	Human sec
c 597	15	78.9	3732	12	ADG15247	cDNA enco	c 670	14.8	77.9	372	3	AAC16429	Human sec
c 598	15	78.9	3732	12	ADf96645	Human PRO	c 671	14.8	77.9	378	3	AAC17871	Human sec
c 599	15	78.9	3732	12	ADG05830	Human PRO	c 672	14.8	77.9	380	13	ACF85487	Human SIR
c 600	15	78.9	3732	12	ADG23414	Novel hum	c 673	14.8	77.9	391	4	AAK63479	Human imm
c 601	15	78.9	3732	12	ADG03703	Human PRO	c 674	14.8	77.9	394	6	ABN21285	Human ORF
c 602	15	78.9	3732	12	ADG24604	Novel hum	c 675	14.8	77.9	400	6	ABN18567	Human ORF

C 676	14.8	77.9	403	10	ABX61350	Abx61250 Arabidops	C 749	14.8	77.9	718	13	ADJ08344	Adj08344 Human car
C 677	14.8	77.9	408	12	ADP93306	Adp93306 Cotton ex	C 750	14.8	77.9	718	13	ADJ08341	Adj08341 Human car
C 678	14.8	77.9	408	13	ADRE1813	Adre1813 Cotton cD	C 751	14.8	77.9	741	10	ADC72575	Adc72575 DNA Seq I
C 679	14.8	77.9	419	3	AACS9229	Aacs9229 Human sec	C 752	14.8	77.9	741	10	ADK57549	Adk57549 Plant DNA
C 680	14.8	77.9	425	5	ADL36537	Adl36537 Human ova	C 753	14.8	77.9	806	4	AAS29770	Aas29770 Human cyt
C 681	14.8	77.9	425	5	ADL71379	Adl71379 Human ova	C 754	14.8	77.9	806	4	ABK43813	Abk43813 DNA encod
C 682	14.8	77.9	430	6	ABX66675	Abx66675 Helicobac	C 755	14.8	77.9	806	12	ADI54200	Adi54200 cDNA enco
C 683	14.8	77.9	430	6	ABX65833	Abx65833 Helicobac	C 756	14.8	77.9	809	13	ADT18340	Adt18340 Plant cDN
C 684	14.8	77.9	438	10	ADBS5258	Adbs5258 Toxicity-	C 757	14.8	77.9	849	6	ABQ68362	Abq68362 Listeria
C 685	14.8	77.9	438	10	ADB49774	Adb49774 Primary r	C 758	14.8	77.9	866	3	AAC50326	Aac50326 Arabidops
C 686	14.8	77.9	444	6	ABL90080	Abl90080 Human pol	C 759	14.8	77.9	868	3	AAC33989	Aac33989 Arabidops
C 687	14.8	77.9	451	9	ACH40960	Ach40960 Human foe	C 760	14.8	77.9	910	13	ADX35279	Adx35279 Plant ful
C 688	14.8	77.9	470	10	ADF82633	Adf82633 Leukaemia	C 761	14.8	77.9	913	8	ABZ36339	Abz36339 Human sec
C 689	14.8	77.9	473	3	AAC07067	Aac07067 Human sec	C 762	14.8	77.9	931	3	AAC79995	Aac79995 Human sec
C 690	14.8	77.9	477	9	ACH25372	Ach25372 Human adu	C 763	14.8	77.9	939	8	ACA27710	Aca27710 Prokaryot
C 691	14.8	77.9	478	3	AAC07066	Aac07066 Human sec	C 764	14.8	77.9	962	10	ADG75799	Adg75799 Human pro
C 692	14.8	77.9	484	4	AAK69537	Aak69537 Human imm	C 765	14.8	77.9	966	2	AAT72290	Aat72290 Rat D cla
C 693	14.8	77.9	486	6	ABL84174	Abl84174 Human ova	C 766	14.8	77.9	966	4	AAC86791	Aac86791 Nucleotid
C 694	14.8	77.9	486	13	ACF88662	Acf88662 Human SIR	C 767	14.8	77.9	971	13	ADR63270	Adr63270 Cotton cD
C 695	14.8	77.9	489	5	ADL42921	Adl42921 Human ova	C 768	14.8	77.9	993	10	ADB69693	Adb69693 C. neofor
C 696	14.8	77.9	497	13	ADR60526	Adr60526 Cotton cD	C 769	14.8	77.9	993	13	ADX31630	Adx31630 Plant ful
C 697	14.8	77.9	502	9	ACH38572	Ach38572 Human end	C 770	14.8	77.9	1001	3	AAC57756	Aac57756 Arachidon
C 698	14.8	77.9	504	3	ACH37174	Aac37174 Arabidops	C 771	14.8	77.9	1001	3	AAC57754	Aac57754 Arachidon
C 699	14.8	77.9	504	14	AEEL13595	Aeel13595 Hamster C	C 772	14.8	77.9	1001	5	AAS80255	Aas80255 DNA encod
C 700	14.8	77.9	504	14	AEEL17237	Aeel17237 Hamster S	C 773	14.8	77.9	1008	4	AAH21004	Aah21004 Bovine-de
C 701	14.8	77.9	515	12	ADL42015	Adl42015 Plant tra	C 774	14.8	77.9	1014	6	ABL34387	AbL34387 Human imm
C 702	14.8	77.9	515	12	ADO02566	Ado02566 Soybean o	C 775	14.8	77.9	1022	11	ACL32732	ACL32732 Rice abio
C 703	14.8	77.9	518	14	AEEL15424	Aeel15424 Hamster S	C 776	14.8	77.9	1059	6	ABK49067	Abk49067 Rice PCR
C 704	14.8	77.9	518	14	AEEL11782	Aeel11782 Hamster c	C 777	14.8	77.9	1059	8	ACF30937	Acf30937 Rice thro
C 705	14.8	77.9	519	2	AAQ25000	Aaq25000 Human int	C 778	14.8	77.9	1059	12	ADI09996	Adi09996 Rice chro
C 706	14.8	77.9	519	2	AAQ24999	Aaq24999 Human int	C 779	14.8	77.9	1076	13	ADI42240	Adi42240 Plant tra
C 707	14.8	77.9	522	1	AAAN70529	Aan70529 Sequence	C 780	14.8	77.9	1082	8	ADZ36166	Adz36166 Human sec
C 708	14.8	77.9	540	13	ADX27695	Adx27695 Plant ful	C 781	14.8	77.9	1083	4	ABK43497	Abk43497 DNA encod
C 709	14.8	77.9	544	10	ADF80589	Adf80589 Leukaemia	C 782	14.8	77.9	1083	12	ADI53884	Adi53884 cDNA enco
C 710	14.8	77.9	551	12	ACH79119	Ach79119 Human gen	C 783	14.8	77.9	1099	11	ADM44828	Adm44828 Insect re
C 711	14.8	77.9	553	4	AAK61206	Aak61206 Human imm	C 784	14.8	77.9	1128	1	AAAG60272	Aan60272 Sequence
C 712	14.8	77.9	576	6	ABN68673	Abn68673 Streptoco	C 785	14.8	77.9	1132	10	ADZ72892	Adz72892 Human end
C 713	14.8	77.9	579	13	ADQ51894	Adq51894 Novel can	C 786	14.8	77.9	1146	10	ADZ72892	Adz72892 C. neofor
C 714	14.8	77.9	585	2	AAK81566	Aax81566 cDNA enco	C 787	14.8	77.9	1155	8	ACA29449	Aca29449 Prokaryot
C 715	14.8	77.9	585	3	AAAG3224	Aaa63224 Human int	C 788	14.8	77.9	1177	2	AAQ79522	Aaq79522 PhoA prom
C 716	14.8	77.9	588	12	ADN49746	Adn49746 Human int	C 789	14.8	77.9	1236	6	ABZ11985	Abz11985 Human pol
C 717	14.8	77.9	588	13	ADU74422	Adu74422 Human int	C 790	14.8	77.9	1236	12	ADM44503	Adm44503 Novel hum
C 718	14.8	77.9	588	14	ADV42752	Adv42752 Human psy	C 791	14.8	77.9	1244	4	AAH229943	Aah229943 S cerevis
C 719	14.8	77.9	593	13	ADQ54714	Adq54714 Novel can	C 792	14.8	77.9	1287	3	ACA28885	Aac42885 Arabidops
C 720	14.8	77.9	596	5	ABV53635	Abv53635 Human pro	C 793	14.8	77.9	1287	8	ADA68530	Ada68530 Arabidops
C 721	14.8	77.9	601	3	AAC39797	Aac39797 Arabidops	C 794	14.8	77.9	1320	11	ADJ31165	Adj31165 Human cDN
C 722	14.8	77.9	602	13	ADU13384	Adul13384 Solid tum	C 795	14.8	77.9	1320	13	ADZ83232	Adz83232 Human lym
C 723	14.8	77.9	610	5	ABA13162	Aba13162 Human ner	C 796	14.8	77.9	1331	12	ADN05627	Adn05627 Antipsori
C 724	14.8	77.9	631	1	AAAN70531	Aan70531 Replicon	C 797	14.8	77.9	1331	14	ADY19279	Ady19279 DNA encod
C 725	14.8	77.9	635	6	ABN63740	Abn63740 Human can	C 798	14.8	77.9	1399	3	AAC50060	Aac50060 Arabidops
C 726	14.8	77.9	637	10	ACD95098	Acd95098 Human col	C 799	14.8	77.9	1402	3	AAC34115	Aac34115 Arabidops
C 727	14.8	77.9	638	10	ADZ72376	Adz72376 DNA Seq I	C 800	14.8	77.9	1452	12	ADK17032	Adk17032 Nanoarcha
C 728	14.8	77.9	663	6	ABQ65935	Abq65935 Arabidops	C 801	14.8	77.9	1456	4	AAI61243	Aai61243 Human int
C 729	14.8	77.9	663	3	AAC38624	Aac38624 Arabidops	C 802	14.8	77.9	1514	11	ADM13728	Adm13728 Human int
C 730	14.8	77.9	663	12	ADM06815	Adm06815 Environme	C 803	14.8	77.9	1537	4	AAI59457	Aai59457 Human pol
C 731	14.8	77.9	676	4	AAAF17557	Aaf17557 Human bre	C 804	14.8	77.9	1583	3	ADZ65262	Adz65262 Human sec
C 732	14.8	77.9	676	4	AAAS46987	Aas46987 Human bre	C 805	14.8	77.9	1583	8	ADA40216	Ada40216 Human sec
C 733	14.8	77.9	676	6	ABK95023	Abk95023 Human bre	C 806	14.8	77.9	1583	10	ADC73743	Adc73743 Human sec
C 734	14.8	77.9	676	6	ABT08642	Abt08642 Human bre	C 807	14.8	77.9	1583	10	ADD37706	Add37706 Human sec
C 735	14.8	77.9	676	6	ABSG3588	Absg3588 Human bre	C 808	14.8	77.9	1583	10	ADL11652	Adel1652 Human sec
C 736	14.8	77.9	676	10	ABT32800	Abt32800 Human tum	C 809	14.8	77.9	1583	10	ADA56376	Ada56376 Gene enco
C 737	14.8	77.9	676	11	ADJ92703	Adj92703 Human bre	C 810	14.8	77.9	1606	6	ABQ72688	Abq72688 Human MDD
C 738	14.8	77.9	676	12	ADA44472	Ada44472 Human cDN	C 811	14.8	77.9	1625	4	AAK51855	Aak51855 Human pol
C 739	14.8	77.9	692	10	ADC72501	Adc72501 DNA Seq I	C 812	14.8	77.9	1708	3	AAC74228	Aac74228 Human sec
C 740	14.8	77.9	692	10	ADK58359	Adk58359 Plant DNA	C 813	14.8	77.9	1720	6	ABS65714	Abs65714 Mouse gen
C 741	14.8	77.9	695	10	ADK56315	Adk56315 Plant DNA	C 814	14.8	77.9	1747	10	ADF81992	Adf81992 Leukaemia
C 742	14.8	77.9	695	10	ADK52756	Adk52756 Plant DNA	C 815	14.8	77.9	1753	4	AAD17359	Aad17359 Human bon
C 743	14.8	77.9	707	6	ABK78003	Abk78003 Bacillus	C 816	14.8	77.9	1758	14	ADZ11518	Adz11518 Mouse cDN
C 744	14.8	77.9	708	6	ABK77843	Abk77843 Bacillus	C 817	14.8	77.9	1773	4	AAD17380	Adad17380 Human bon
C 745	14.8	77.9	718	4	AAAS36229	Aas36229 Human car	C 818	14.8	77.9	1797	14	ADY20644	Ady20644 DNA encod
C 746	14.8	77.9	718	4	AAAS36232	Aas36232 Human car	C 819	14.8	77.9	1803	10	ADP69667	Adp69667 Chimeric
C 747	14.8	77.9	718	10	ADA64926	Ade46926 Human car	C 820	14.8	77.9	1848	14	ADW18554	Adw18554 Pinus rad
C 748	14.8	77.9	718	10	ADZ46923	Ade46923 Human car	C 821	14.8	77.9	1848	14	ADY52945	Ady52945 Tomato ke

C 822	14.8	77.9	1848	14	ADY52395	Novel ket
C 823	14.8	77.9	1848	14	ADY51392	Synechoco
C 824	14.8	77.9	1882	13	ADQ82277	Chicken D
C 825	14.8	77.9	1884	10	ADF14272	Human end
C 826	14.8	77.9	1926	6	ABK35700	cDNA sequ
C 827	14.8	77.9	1934	1	AAAN60201	Interfero
C 828	14.8	77.9	1968	9	ADB10247	Allioiococ
C 829	14.8	77.9	1980	6	ABL49742	Human int
C 830	14.8	77.9	2000	6	ABZ15821	Arabidops
C 831	14.8	77.9	2000	10	ADC08409	Rice DNA
C 832	14.8	77.9	2000	10	ACC61786	Gene sequ
C 833	14.8	77.9	2000	10	ADK63941	Disease t
C 834	14.8	77.9	2025	4	RAK52767	Human pol
C 835	14.8	77.9	2089	4	AAH13958	Human CDN
C 836	14.8	77.9	2091	8	ACA30588	Prokaryot
C 837	14.8	77.9	2093	10	ABV77193	Nucleotid
C 838	14.8	77.9	2100	14	ADZ00538	Human Adl
C 839	14.8	77.9	2102	5	ABA18997	Human ner
C 840	14.8	77.9	2102	5	ABA18996	Human ner
C 841	14.8	77.9	2102	11	ADI31038	Human CDN
C 842	14.8	77.9	2102	13	ADS83105	Human lym
C 843	14.8	77.9	2108	10	ABV77195	Nucleotid
C 844	14.8	77.9	2124	4	AAE81745	Human mem
C 845	14.8	77.9	2151	10	AEF78381	Soybean s
C 846	14.8	77.9	2151	15	AEF94244	Stress re
C 847	14.8	77.9	2155	6	ABL90322	Human pol
C 848	14.8	77.9	2158	13	ADRO7484	Human pol
C 849	14.8	77.9	2187	4	AAH16070	Human CDN
C 850	14.8	77.9	2187	13	ACN400773	Tumour-as
C 851	14.8	77.9	2188	9	ABZ59658	Human sec
C 852	14.8	77.9	2190	5	ABV30186	Human pro
C 853	14.8	77.9	2210	10	ABV77189	Nucleotid
C 854	14.8	77.9	2210	13	ACN37570	Tumour-as
C 855	14.8	77.9	2268	9	RAE57478	Fruit fly
C 856	14.8	77.9	2313	10	ADC91107	E. faeciu
C 857	14.8	77.9	2375	6	ABK89447	Human zin
C 858	14.8	77.9	2388	10	AAAL54532	Tomato pl
C 859	14.8	77.9	2457	10	ADA52870	Human cod
C 860	14.8	77.9	2485	10	ADF14285	Human end
C 861	14.8	77.9	2524	5	ADL62864	Human ova
C 862	14.8	77.9	2535	10	AAAD51682	Human nuc
C 863	14.8	77.9	2616	8	ACA30186	Prokaryot
C 864	14.8	77.9	2760	10	ADE99800	Bacterial
C 865	14.8	77.9	2833	13	ADR14644	Human NF-
C 866	14.8	77.9	2856	14	ADZ14753	ORF DNA e
C 867	14.8	77.9	2856	14	ADZ14489	DNA encod
C 868	14.8	77.9	2889	10	ADK67246	Human can
C 869	14.8	77.9	2889	10	ADK67248	Human can
C 870	14.8	77.9	2934	2	AAAX03030	Human IL-
C 871	14.8	77.9	2946	13	ADT15247	Plant CDN
C 872	14.8	77.9	3086	13	ADK63853	Plant ful
C 873	14.8	77.9	3089	6	ABN97192	Gene #369
C 874	14.8	77.9	3089	10	ADD18665	Human dis
C 875	14.8	77.9	3089	10	ACA56732	Human sig
C 876	14.8	77.9	3089	12	ADL56528	Human pol
C 877	14.8	77.9	3105	13	ADU05460	DNA encod
C 878	14.8	77.9	3111	3	AAAS3888	Beta, Bet
C 879	14.8	77.9	3131	4	AAH57491	Human liv
C 880	14.8	77.9	3146	10	ADB68971	C. neofo
C 881	14.8	77.9	3175	6	ABQ70847	Listeria
C 882	14.8	77.9	3534	8	ACA28380	Prokaryot
C 883	14.8	77.9	3551	11	ADM01925	Human CDN
C 884	14.8	77.9	3551	14	AEC84855	Human CDN
C 885	14.8	77.9	3636	4	AAAS1642	Staphyloc
C 886	14.8	77.9	3816	11	ADM01977	Human CDN
C 887	14.8	77.9	3816	14	AEC84907	Human CDN
C 888	14.8	77.9	3821	12	ADQ63572	Novel hum
C 889	14.8	77.9	3830	14	ADZ49688	Ineulin s
C 890	14.8	77.9	3978	14	ADZ14671	DNA encod
C 891	14.8	77.9	4143	13	ADS46674	Bacterial
C 892	14.8	77.9	4172	13	ADRO7850	Full leng
C 893	14.8	77.9	4224	13	ADS46785	Bacterial
C 894	14.8	77.9	4260	15	AEF78205	Yeast Ede
AAS54749	14.8	77.9	4449	4	AAS54749	Staphyloc
ABL27892	14.8	77.9	4463	4	ABL27892	Drosophil
AAS57190	14.8	77.9	4463	4	AAS57190	DNA encod
ADC35926	14.8	77.9	4463	10	ADC35926	Drosophil
ADR08441	14.8	77.9	4666	13	ADR08441	Full leng
AAX81563	14.8	77.9	5322	2	AAX81563	Cytokine-
AAA63225	14.8	77.9	5322	3	AAA63225	Plasmid V
AAX81562	14.8	77.9	5480	2	AAX81562	Cytokine-
ABL11784	14.8	77.9	5945	4	ABL11784	Drosophil
AAS29815	14.8	77.9	6095	4	AAS29815	Human cyt
ABL10424	14.8	77.9	6225	4	ABL10424	Drosophil
ABL06640	14.8	77.9	6228	4	ABL06640	Drosophil
ADN02625	14.8	77.9	6497	12	ADN02625	Liver dis
ABL12620	14.8	77.9	6580	4	ABL12620	Drosophil
ADL22645	14.8	77.9	6597	11	ADL22645	Human dis
ABL21538	14.8	77.9	6968	4	ABL21538	Drosophil
ADC66457	14.8	77.9	7174	10	ADC66457	Human enz
AAS32738	14.8	77.9	7549	4	AAS32738	Human gen
AAS32737	14.8	77.9	7549	4	AAS32737	Human gen
AAL54534	14.8	77.9	7650	10	AAL54534	Tomato pl
AAX89439	14.8	77.9	7680	3	AAX89439	14-3-3 si
AAL54535	14.8	77.9	7717	10	AAL54535	Tomato pl
AAL54533	14.8	77.9	7998	10	AAL54533	Tomato pl
AAK65546	14.8	77.9	8223	4	AAK65546	Human imm
ABL06304	14.8	77.9	9012	4	ABL06304	Drosophil
ABT06535	14.8	77.9	10034	6	ABT06535	Human 14-
ADM83712	14.8	77.9	10034	11	ADM83712	Human cyc
AEC22596	14.8	77.9	10034	14	AEC22596	Human 14-
AED18633	14.8	77.9	10034	14	AED18633	Fibrotic
ADN37124	14.8	77.9	10410	12	ADN37124	DNA sequ
AAS41739	14.8	77.9	11022	4	AAS41739	Genomic s
AAK80617	14.8	77.9	11022	4	AAK80617	Human imm
AAK48974	14.8	77.9	11220	3	AAK48974	Arabidops
ABL15550	14.8	77.9	12138	4	ABL15550	Drosophil
ABL15542	14.8	77.9	12161	4	ABL15542	Drosophil
ABL20646	14.8	77.9	12413	4	ABL20646	Drosophil
AAAL05001	14.8	77.9	12970	4	AAAL05001	Human rep
ABL97894	14.8	77.9	12970	4	ABL97894	Human tes
ADG20558	14.8	77.9	15620	12	ADG20558	Human RNA
AAS41741	14.8	77.9	16291	4	AAS41741	Genomic s
AAK80619	14.8	77.9	16291	4	AAK80619	Human imm
ABD33054	14.8	77.9	22235	13	ABD33054	Human can
AEA61257	14.8	77.9	25049	14	AEA61257	Human PLA
AEA61202	14.8	77.9	25049	14	AEA61202	Human PLA
ADA02987	14.8	77.9	27587	9	ADA02987	Mouse Cd2
ADB72725	14.8	77.9	27587	10	ADB72725	Mouse Cd2
ADC85467	14.8	77.9	27587	10	ADC85467	Mouse Cd2
ADM74582	14.8	77.9	27587	12	ADM74582	Murine ca
ABX11387	14.8	77.9	29921	8	ABX11387	DNA encod
AAS29814	14.8	77.9	29973	4	AAS29814	Human cyt
ADJ12408	14.8	77.9	29973	12	ADJ12408	Human fragm
AAS29831	14.8	77.9	32205	4	AAS29831	Human cyt
AAS33426	14.8	77.9	32205	4	AAS33426	DNA encod
AAAL04789	14.8	77.9	32249	4	AAAL04789	Human rep
ABL97684	14.8	77.9	32249	4	ABL97684	Human tes
ABK50980	14.8	77.9	36568	6	ABK50980	Human sol
AAK85853	14.8	77.9	39989	4	AAK85853	Human imm
AEA61140	14.8	77.9	40085	14	AEA61140	Human FMO
ACN44568	14.8	77.9	43799	11	ACN44568	Mouse gen
ADZ12720	14.8	77.9	45239	14	ADZ12720	Murine ca
AAAL12437	14.8	77.9	45000	4	AAAL12437	DNA encod
ABL04952	14.8	77.9	52872	4	ABL04952	Rice cult
ACF30939	14.8	77.9	53905	8	ACF30939	Rice cult
ADIO9998	14.8	77.9	53905	12	ADIO9998	Rice cult
ADV66784	14.8	77.9	53905	14	ADV66784	Rice Chro
AAK80620	14.8	77.9	59560	4	AAK80620	Human imm
AAK71550	14.8	77.9	59560	4	AAK71550	Human imm
ACN44724	14.8	77.9	59589	11	ACN44724	Mouse gen
ADR72312	14.8	77.9	63045	13	ADR72312	A truncat
ADE43742	14.8	77.9	63824	10	ADE43742	Human KNS
ADH54220	14.8	77.9	63824	10	ADH54220	Polymorph
ADH54221	14.8	77.9	63824	12	ADH54221	Human KNS
ADH54221	14.8	77.9	63824	12	ADH54221	Human KNS

c 968 14.8 77.9 65558 14 AEA61193
969 14.8 77.9 66126 13 ABD33341 Murine ca
970 14.8 77.9 72821 13 ABD32900 Mouse can
971 14.8 77.9 81679 12 ADQ59189 MSI-H car
972 14.8 77.9 82660 11 ACN45192 Mouse gen
973 14.8 77.9 82952 6 ABN85766 Arabidops
974 14.8 77.9 83400 12 ADP07906 Human RAD
975 14.8 77.9 90885 12 ADK16049_4 Continuation (5 of
976 14.8 77.9 92407 4 AAF28549 Genomic f
977 14.8 77.9 94330 11 ACN44662 Human gen
978 14.8 77.9 96593 9 ADA02750 Human IRF
979 14.8 77.9 96593 10 ADB72488 Human IRP
980 14.8 77.9 96593 12 ADC85230 Human IRF
981 14.8 77.9 96593 12 ADM74345 Human car
982 14.8 77.9 96593 9 ADA02936 Human BRA
983 14.8 77.9 96595 10 ADB72674 Human BRA
984 14.8 77.9 96595 10 ADC85416 Human Bra
985 14.8 77.9 96595 12 ADM74531 Human car
986 14.8 77.9 96596 10 ADC85287 Mouse Fis
987 14.8 77.9 96597 9 ADA02807 Mouse Fis
988 14.8 77.9 96597 10 ADB72545 Mouse Fis
989 14.8 77.9 96597 12 ADM74402 Murine ca
990 14.8 77.9 96599 9 ADA02831 Mouse Mef
991 14.8 77.9 96599 10 ADB72569 Mouse Mef
992 14.8 77.9 96599 10 ADC85310 Human Pou
993 14.8 77.9 96599 12 ADM74426 Murine ca
994 14.8 77.9 99916 6 ADI03931 Human enz
995 14.8 77.9 101505 11 ACN44694 Human gen
996 14.8 77.9 101786 3 AAF22293 BAC conta
997 14.8 77.9 110000 2 AAT42063_17 Continuation (18 o
998 14.8 77.9 110000 2 AAV21209_11 Continuation (12 o
999 14.8 77.9 110000 2 AAX20248_01 Continuation (2 of
1000 14.8 77.9 110000 2 AAZ01425_01 Continuation (2 of

ALIGNMENTS

RESULT 1
ADI28529
ID ADI28529 standard; DNA, 19 BP.
XX AC ADI28529;
XX AC
XX 22-APR-2004 (first entry)
XX DT
XX DE Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA primer.
XX XX
KW Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
KW broncholytic; gene therapy; PCR; primer; ss.
XX XX
OS Homo sapiens.
XX XX
XX WO2004001060-A2.
XX XX
XX 31-DEC-2003.
XX XX
XX 20-JUN-2003; 2003WO-US019255.
XX XX
XX 20-JUN-2002; 2002US-0390850P.
XX PR 29-AUG-2002; 2002US-0407006P.
XX XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX XX
XX Whitney GS, Opitck G, Garulacan L, Ramanathan CS, Mckinnon M;
XX Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX WPI; 2004-090973/09.
XX XX
XX New nucleic acid molecule encoding a human G-protein coupled receptor.
XX RAI3, useful for preventing, treating or ameliorating chronic obstructive
XX pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms

PT of COPD.
XX Disclosure; SEQ ID NO 32; 301pp; English.
PS The present sequence is that of an antisense PCR primer based on DNA
XX encoding a peptide ADI28458 from human G-protein coupled receptor
CC retinoic acid induced 3 (RAI-3) ADI28460. The peptide was identified in
CC studies of RAI-3 phosphorylation in response to cigarette smoke. The
CC primer can be used in a PCR to identify individuals at risk for
CC developing chronic obstructive pulmonary disease (COPD). RAI-3 is
CC tyrosine phosphorylated and/or is associated/complexed with tyrosine
CC phosphorylated proteins only in cells that have been exposed to cigarette
CC smoke. Since RAI-3 is primarily expressed in lung tissue, and since a
CC cigarette smoke is a major causative factor of COPD, RAI-3 provides a
CC novel cellular target for identifying modulators, e.g. agonists or
CC antagonists, useful for the treatment and/or prevention of COPD and
CC related disorders such as emphysema and chronic bronchitis. RAI-3
CC modulators, e.g. agonists and antagonists, especially antisense
CC compounds, can be used to treat COPD and other disorders an diseases
CC associated with regulation of NF-kB and/or its associated or interacting
CC signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
CC the RAI-3 gene are useful for determining COPD association in
CC individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
CC preventing, treating or ameliorating disorders related to aberrant GPCR
CC signaling or cell cycle regulation, pulmonary disorders, inflammatory
CC lung disorders, COPD, the underlying symptoms of COPD, COPD-related
CC disorders or conditions, autoimmune disorders, disorders related to
CC hyperimmune activity, inflammatory conditions, disorders related to
CC aberrant acute phase responses, hypercongenital conditions, birth
CC defects, necrotic lesions, wounds, organ transplant rejection, renal
CC diseases, ischaemia-reperfusion injury, heart disorders, disorders
CC related to aberrant signal transduction, proliferation disorders,
CC cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
CC ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
CC nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
CC rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
CC pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
XX regulation (all claimed).
XX XX
SQ Sequence 19 BP; 3 A; 4 C; 2 G; 10 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. NO. 40;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGCTTTG 19
Db 1 CTTTACTTCATAGCTTTG 19
RESULT 2
ADT96418/c
ID ADT96418 standard; cDNA; 497 BP.
XX AC ADT96418;
XX XX
XX 16-DEC-2004 (first entry)
XX DT
XX DE Colon cancer associated human cDNA sequence #1925.
XX XX
XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
XX humoral immune response; cellular immune response; cytostatic;
XX immunostimulant; human; ss.
XX XX
XX Homo sapiens.
XX OS
XX US2003087818-A1.
XX FN
XX 08-MAY-2003.
XX PD
XX

```
PF 01-FEB-2002; 2002US-00066543.
XX
XX
XX 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1937; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
XX Sequence 497 BP; 126 A; 130 C; 129 G; 112 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 19; DB 11; Length 497;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGTCCTTTG 19
Db 385 CTTTACTTCATAGTCCTTTG 367

RESULT 3
ADX42900/c
ID ADX42900 standard; cDNA; 497 BP.
XX
XX ADX42900;
XX
XX 21-APR-2005 (first entry)
XX
XX Human cDNA encoding colon cancer protein SEQ ID NO 1937.
XX
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
XX ss; gene.
XX
XX Homo sapiens.
XX
XX WO200274156-A2.
PN

XX 26-SEP-2002.
XX
XX 01-FEB-2002; 2002WO-US002870.
XX
XX 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1937; 244pp; English.
XX
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
XX Sequence 497 BP; 126 A; 130 C; 129 G; 112 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 19; DB 11; Length 497;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGTCCTTTG 19
Db 385 CTTTACTTCATAGTCCTTTG 367

RESULT 4
ADT96292
ID ADT96292 standard; cDNA; 552 BP.
XX
XX ADT96292;
XX
XX 16-DEC-2004 (first entry)
XX
XX Colon cancer associated human cDNA sequence #1799.
XX
XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
XX humoral immune response; cellular immune response; cytostatic;
XX immunostimulant; human; ss.
XX
XX Homo sapiens.
XX
XX US2003087818-A1.
XX
XX 08-MAY-2003.
XX
XX 01-FEB-2002; 2002US-00066543.
XX
XX 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
```



```
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Chenault RA, Xu J, Indirias CV, Lodes MJ, Secrist H;
XX PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX DR WPI; 2003-040540/03.
XX
XX PT New isolated nucleic acids and polypeptides capable of eliciting a
XX PT humoral and/or cellular immune response, useful for diagnosing,
XX PT preventing or treating cancer, particularly colon cancer.
XX PS Claim 1; SEQ ID NO 1811; 87pp; English.
XX
XX CC The invention relates to polynucleotide and polypeptide sequences
XX CC associated with cancer, particularly colon cancer. Also disclosed are (i)
XX CC an expression vector comprising the polynucleotide, (ii) a host cell
XX CC transformed or transfected with the expression vector, (iii) an isolated
XX CC antibody, or its antigen-binding fragment, which specifically binds to
XX CC the polypeptide, (iv) a method of detecting or determining the presence
XX CC of cancer in a patient, (v) a fusion protein comprising at least one of
XX CC the polypeptides, (vi) an oligonucleotide that hybridizes to the
XX CC polynucleotide sequence under highly stringent conditions, and (vii) a
XX CC method of stimulating and/or expanding T cells specific for a tumour
XX CC protein. The polypeptide specifically comprises the amino acid sequence
XX CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
XX CC frames (ORFs) of C636S. These polypeptides are encoded by the
XX CC polynucleotide sequences, where both are capable of eliciting a humoral
XX CC and/or cellular immune response. The polynucleotides, polypeptides, and
XX CC antibodies are useful for diagnosing, preventing or treating cancer,
XX CC particularly colon cancer. The polynucleotide and polypeptide sequences
XX CC are also useful in DNA strand invasion, antisense inhibition, mutational
XX CC analysis, nucleic acid purification, isolation of transcriptionally
XX CC active genes, blocking or transcription factor binding, genome cleavage
XX CC or in situ hybridization, and as enhancers of transcription or
XX CC biomarkers. This sequence represents a human colon cancer associated
XX CC cDNA. Note: The sequence data for this patent was obtained in electronic
XX CC format directly from the USPTO web site at seqdata.uspto.gov
XX
XX SQ Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 113 CTTTACTTCATAGTCTTTG 131

RESULT 5
ADX42774
ID ADX42774 standard; cDNA; 552 BP.
XX AC ADX42774;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE Human cDNA encoding colon cancer protein SEQ ID NO 1811.
XX
XX KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
XX KW ss; gene.
XX
XX OS Homo sapiens.
XX
XX PN WO200274156-A2.
XX
XX PD 26-SEP-2002.
XX
XX PX 01-FEB-2002; 2002WO-US002870.
XX
XX PR 02-FEB-2001; 2001US-0267400P.
XX PR 07-FEB-2001; 2001US-0267382P.

PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Jiang Y, Chenault RA, Xu J, Indirias CV, Lodes MJ, Secrist H;
XX PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX XX WPI; 2003-040540/03.
XX
XX PT New isolated nucleic acids and polypeptides capable of eliciting a
XX PT humoral and/or cellular immune response, useful for diagnosing,
XX PT preventing or treating cancer, particularly colon cancer.
XX PS Claim 1; SEQ ID NO 1811; 244pp; English.
XX
XX CC The invention relates to a new isolated nucleic acid. The nucleic acids,
XX CC polypeptides, antibodies are useful for diagnosing, preventing or
XX CC treating cancer, particularly colon cancer. The nucleic acid and
XX CC polypeptides are also useful in DNA strand invasion, antisense
XX CC inhibition, mutational analysis, nucleic acid purification, isolation of
XX CC transcriptionally active genes, blocking or transcription factor binding,
XX CC genome cleavage or in situ hybridization, and as enhancers of
XX CC transcription or biomarkers. The kits are useful for detecting antibody
XX CC binding. The present sequence represents a human cDNA encoding a colon
XX CC cancer protein.
XX
XX SQ Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 113 CTTTACTTCATAGTCTTTG 131

RESULT 6
AAZ90046/c
ID AAZ90046 standard; cDNA; 603 BP.
XX AC AAZ90046;
XX
XX DT 09-MAY-2000 (first entry)
XX
XX DE Hydrophobic domain containing protein clone HP10549 coding sequence.
XX
XX KW Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
XX KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;
XX KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
XX KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
XX KW anti-inflammatory; infection; bodily characteristic; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200000506-A2.
XX
XX PD 06-JAN-2000.
XX
XX PF 18-JUN-1999; 99WO-JP003242.
XX
XX PR 26-JUN-1998; 98JP-00180008.
XX
XX PA (SAGA ) SAGAMI CHEM RES CENT.
XX PA (PROT-) PROTEGENE INC.
XX
XX PI Kato S, Kimura T;
XX
XX DR WPI; 2000-160665/14.
XX DR P-PSDB; AAY78809.
XX
```


PT Novel human proteins having hydrophobic domains used for research and
XX diagnostic purposes.
XX
PS Claim 3; Page 88; 117pp; English.
XX
CC This sequence represents the hydrophobic domain containing protein, clone
CC HP10549 coding region. The sequence is isolated from a human stomach
CC cancer cell line. The invention relates to human proteins with
CC hydrophobic domains, the DNA and the cDNA encoding them. The
CC polynucleotides and proteins are predicted to have biological activities
CC which make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals. Suggested activities include
CC nutritional activity (nutritional source or supplement); cytokine and
CC cell proliferation/differentiation activity; immune stimulating (e.g. as
CC vaccines) or suppressing activity (e.g. to treat various immune
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC organ transplantation); haematopoiesis regulating activity (e.g. in
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or cardiac cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
SQ Sequence 603 BP; 133 A; 179 C; 137 G; 154 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 3; Length 603;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGCTCTTG 19
Db 591 CTTTACTTCATAGCTCTTG 573
RESULT 7
ABQ58527
ID ABQ58527 standard; cDNA; 620 BP.
XX
AC ABQ58527;
XX
XX 02-AUG-2002 (first entry)
XX
XX Human colon cancer related nucleotide sequence SEQ ID NO:2222.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200229086-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US030732.
XX
XX 02-OCT-2000; 2000US-0237271P.
XX
XX (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX Thiagalingam A, Lewis ME;
XX

XX WPI; 2002-426115/45.
DR
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 620 BP; 152 A; 154 C; 161 G; 142 T; 0 U; 11 Other;
Query Match 100.0%; Score 19; DB 6; Length 620;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGCTCTTG 19
Db 113 CTTTACTTCATAGCTCTTG 131
RESULT 8
ABQ59698
ID ABQ59698 standard; cDNA; 634 BP.
XX
XX AC ABQ59698;
XX
XX 02-AUG-2002 (first entry)
XX
XX Human colon cancer related nucleotide sequence SEQ ID NO:3393.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200229086-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US030732.
XX
XX 02-OCT-2000; 2000US-0237271P.
XX
XX (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX Thiagalingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX

CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists

XX
SQ Sequence 634 BP; 153 A; 154 C; 168 G; 146 T; 0 U; 13 Other;
Query Match 100.0%; Score 19; DB 6; Length 634;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 117 CTTTACTTCATAGTCTTTG 135

RESULT 9
ID ABT22328 standard; DNA; 642 BP.
XX AC ABT22328;
XX DT 16-APR-2003 (first entry)
XX DE Breast cancer marker gene SEQ ID No 701.
XX KW Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;
KW chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;
KW surrogate marker gene; pharmacodynamic marker gene; transgenic animal;
KW human; ds.
XX OS Homo sapiens.
XX PN WO200285298-A2.
XX PD 31-OCT-2002.
XX PF 19-APR-2002; 2002WO-US012612.
XX PR 20-APR-2001; 2001US-0285163P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;
XX WPI; 2003-093053/08.
XX PT Novel isolated polypeptide encoded by breast cancer marker gene, useful
PT for diagnosing, staging, monitoring, prognosing and treating diseases
PT associated with breast cancer.
XX PS Disclosure; Page 187-188; 725pp; English.
XX CC The invention relates to an isolated polypeptide encoded by a breast
CC cancer marker gene comprising any of 1417 21-805 nucleotide sequences,
CC given in the specification. The methods of the invention are useful for
CC diagnosing patients having an identified breast mass or symptoms
CC associated with breast cancer, to diagnose breast cancer or its
CC precursors, and for monitoring the efficacy of treatment of a breast
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also
CC useful for evaluating a patient before, after or during therapy, to

CC evaluate the reduction in a tumour burden. The breast cancer marker gene
CC proteins are useful as immunogens for raising antibodies, by immunising a
CC mammal with a breast cancer marker protein. The marker proteins are
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to
CC identify other proteins which bind to or interact with the marker
CC proteins. The breast cancer marker genes are useful as surrogate marker
CC genes for one or more disorders, disease states or conditions leading to
CC disease states, in particular, breast cancers. The breast cancer marker
CC genes are useful as pharmacodynamic marker genes. An antibody which
CC selectively binds to a protein of a breast cancer marker gene is useful
CC for treating cancers, particularly breast cancers. The host cell of the
CC invention is useful for producing non-human transgenic animals. This
CC polynucleotide sequence represents one of the breast cancer marker genes
CC of the invention

XX
SQ Sequence 642 BP; 148 A; 160 C; 180 G; 142 T; 0 U; 12 Other;
Query Match 100.0%; Score 19; DB 10; Length 642;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 143 CTTTACTTCATAGTCTTTG 161

RESULT 10
ADO30035/c
ID ADO30035 standard; cDNA; 1074 BP.
XX AC ADO30035;
XX DT 29-JUL-2004 (first entry)
XX DE Human GPCR RAI3 polynucleotide, SEQ ID NO:1137.
XX KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cyostatic; antiinflammatory; vasotropic; antiarrhythmic;
KW CNS; central nervous system; respiratory; antiarrhythmic;
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; Gene therapy; GPCR modulator; human;
KW gene; ss.
XX OS Homo sapiens.
XX PN WO2004040000-A2.
XX PD 13-MAY-2004.
XX PF 09-SEP-2003; 2003WO-US028226.
XX PR 09-SEP-2002; 2002US-0409303P.
XX PR 09-APR-2003; 2003US-0461329P.
XX PA (PRIM-) PRIMAL INC.
XX PI Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX WPI; 2004-390329/36.
XX DR P-PSDB; ADO29632.
XX PT Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition

associated with GPCR dysfunction e.g. autoimmune diseases, angina pectoris, Parkinson's disease.

Claim 151; SEQ ID NO 1137; 542pp; English.

The invention relates to human and mouse G protein-coupled receptors (GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several mice, each of which has a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridise to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia); disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders; blood disorders (e.g., anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and thyroid (e.g., cancers). The present sequence represents a GPCR-encoding nucleic acid of the invention. Note: The full sequence data for this patent did not form part of the printed specification; those sequences not shown were obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1074 BP; 215 A; 316 C; 259 G; 284 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 1074;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
|||||
Db 1059 CTTTACTTCATAGCTTTG 1041
|||||

RESULT 11
AAZ61776/c
ID AAZ61776 standard; cDNA; 1212 BP.
XX AAC61776;
XX 27-MAR-2000 (first entry)
XX cDNA encoding human skin cell transmembrane protein, SEQ ID NO:249.
KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumour vascularisation; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX Homo sapiens.
XX WO9955865-A1.
XX 04-NOV-1999.
XX 29-APR-1999; 99WO-NZ000051.
XX

29-APR-1998; 98US-00069726.
09-NOV-1998; 98US-00188930.
(GENE-) GENESIS RES & DEV CORP LTD.
Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
WPI; 2000-072177/06.
Novel polynucleotides useful for the treatment of various conditions including wounds and cancer.
Claim 1; Page 155; 235pp; English.
The invention relates to novel nucleic acid sequences derived from rat dermal papilla, human keratinocytes and neonatal foreskin fibroblasts, and mouse embryonic skin, keratinocyte stem cells and transit amplifying cells. Polypeptides of the invention may be used to treat inflammation, cancer and neurological diseases. The proteins may be used to stimulate the growth and motility of keratinocytes, to inhibit the growth of cancer cells, to modulate angiogenesis and tumour vascularisation, to modulate skin inflammation, to modulate epithelial cell growth and to inhibit binding of HIV-1 to leukocytes. The invention may also be used to treat growth and developmental defects, skin wounds and hair follicle disorders. Sequences AAZ61606-Z61832 represent cDNA sequences derived from several mouse, rat or human skin cell types. Sequences AAZ61606-Z61649, AAZ61725-Z61765, AAZ61802-Z61811 and AAZ61826 encode proteins with an N-terminal signal sequence, indicating that the proteins are secreted. Sequences AAZ61650-Z61668, AAZ61766-Z61780, AAZ61812-Z61817 and AAZ61827-Z61829 encode proteins with one or more putative transmembrane domains
Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
|||||
Db 1031 CTTTACTTCATAGCTTTG 1013
|||||

RESULT 12
AAC99709/c
ID AAC99709 standard; cDNA; 1212 BP.
XX AAC99709;
XX 08-MAR-2001 (first entry)
XX Skin cell cDNA, SEQ ID NO: 249.
KW Human; skin cell; cytostatic; antiinflammatory; anti-HIV; neurotropic;
KW neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease; ss.
XX Homo sapiens.
XX WO200069884-A2.
XX 23-NOV-2000.
XX 15-MAY-2000; 2000WO-NZ000075.
XX 14-MAY-1999; 99US-00312283.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;
XX Murison JG;

DR WPI; 2001-007495/01.
DR P-PSDB; AAB55958.
XX
PT New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases.
XX
XX Claim 1; Page 210-211; 352pp; English.
PS
SS The present polynucleotide encodes a polypeptide which is expressed in
CC mammalian skin cells. The polypeptide is useful for stimulating
CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC tumours, modulating skin inflammation, stimulating the growth of
CC epithelial cells, inhibiting the binding of human immunodeficiency virus
CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
CC neurological diseases. The polynucleotide can be used as a marker, in the
CC identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns
XX
XX Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 19; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 1031 CTTTACTTCATAGTCTTTG 1013
RESULT 13
ABL34861/c
ID ABL34861 standard; cDNA; 1212 BP.
XX
XX ABL34861;
AC
XX
XX 04-APR-2002 (first entry)
XX
XX Human cDNA isolated from skin cells SEQ ID NO: 249.
DE
XX
XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulneryary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;
KW ss.
XX
XX Homo sapiens.
OS
XX
XX WO200190357-A1.
FN
XX
XX 29-NOV-2001.
PD
XX
XX 24-MAY-2001; 2001WO-NZ000099.
PF
XX
XX 24-MAY-2000; 2000US-0206650P.
PR
XX
XX 25-JUL-2000; 2000US-0221232P.
PR
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA
XX
XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;
PI Kumble KD;
PI
XX
XX WPI; 2002-122020/16.
DR
XX
XX New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.
XX
XX Claim 1; Page 175; 466pp; English.
PS
SS The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences

CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC cDNA of the invention
XX
XX Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 19; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 1031 CTTTACTTCATAGTCTTTG 1013
RESULT 14
ACA04775/c
ID ACA04775 standard; cDNA; 1228 BP.
XX
XX ACA04775;
AC
XX
XX 28-MAY-2003 (first entry)
XX
XX cDNA encoding human membrane associated protein fragment #223.
DE
XX
XX Human; ss; gene; microarray; membrane-associated protein; neuropathology;
KW immunopathology; pancreatic disease; cancer; diabetes; hyperlipidaemia;
KW pancreatic cholera; Alzheimer's disease; Huntington's disease; sarcoma;
KW fibrocystic disease; leukaemia; adenocarcinoma; AIDS; allergy; anaemia;
KW asthma; gout; dementia.
XX
XX Homo sapiens.
OS
XX
XX US6492505-B1.
FN
XX
XX 10-DEC-2002.
PD
XX
XX 31-JAN-2000; 2000US-00495050.
PF
XX
XX 01-FEB-1999; 99US-0118318P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Reddy R, Guegler KJ, Au-Young J;
PI
XX
XX WPI; 2003-327324/31.
DR
XX
XX Combination for research/diagnostic applications and for monitoring
PT treatment of e.g., cancer, comprises polynucleotides comprising a
PT fragment of gene encoding membrane-associated proteins, receptors or ion
PT channels.
XX
XX Claim 1; Col 215-216; 147pp; English.
PS
XX
XX The invention relates to a combination comprising several polynucleotide
CC sequences comprising a fragment of gene encoding membrane-associated
CC proteins, receptors or ion channels. The combination is useful as a
CC probe, for research and diagnostic applications, for monitoring the
CC expression of several expressed polynucleotides, in the diagnosis and
CC monitoring of treatment of pancreatic disease, cancer, immunopathology or
CC neuropathology, for investigating an individual's predisposition to the
CC above disease, in genetic or gene expression analysis of polynucleotide
CC sequences, to investigate cellular responses to infection or drug
CC treatment, as hybridisable array elements in a microarray, to purify a
CC subpopulation of mRNAs, cDNAs or genomic fragments in a sample, in
CC diagnostics, prognostics and treatment regimens, in drug discovery and
CC development, in toxicological and carcinogenicity studies, and in
CC forensics or pharmacogenomics, to monitor the progression of disease, to
CC monitor the efficacy of treatment, to diagnose the conditions of the

CC pancreas e.g. diabetes, pancreatic cholera, hyperlipidaemia or
 CC fibrocystic disease, to diagnose a cancer e.g. leukaemia, adenocarcinoma
 CC or sarcoma, to diagnose immunopathologies e.g. AIDS, allergies, anaemia,
 CC asthma or gout, to diagnose neuropathologies e.g. Alzheimer's disease,
 CC dementia or Huntington's disease, to rapidly screen large numbers of
 CC candidate drug molecules and as query sequences against GenBank,
 CC SwissProt, BLOCKS and PRINTS databases. The combination is employed to
 CC fine tune the treatment regimen and thus the expression patterns
 CC associated with undesirable side effects are avoided. The present
 CC sequence represents a cDNA encoding a fragment of gene encoding human
 CC membrane-associated proteins, receptors or ion channels
 XX
 SQ Sequence 1228 BP; 272 A; 341 C; 283 G; 332 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 8; Length 1228;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19

Db 582 CTTTACTTCATAGTCTTTG 564

RESULT 15

AEA00112/c

ID AEA00112 standard; cDNA; 1460 BP.

XX AEA00112;

AC AEA00112;

DT 28-JUL-2005 (first entry)

XX Human TAT143 cDNA SEQ ID NO:64.

XX ss; gene; tumor-associated antigen; cytostatic; breast tumor;
 KW endocrine disease; gynecology and obstetrics; neoplasm; colon tumor;
 KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
 KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
 KW ovary tumor; skin tumor; liver tumor.

XX Homo sapiens.

XX Key Location/Qualifiers

FH 128..1201

FT /*tag= a

FT /product= "TAT143"

XX US2005106644-A1.

XX 19-MAY-2005.

XX 08-SEP-2004; 2004US-00936626.

XX 20-JUN-2001; 2001US-0299500P.

XX 29-JUN-2001; 2001US-0301880P.

XX 18-SEP-2001; 2001US-0323268P.

XX 19-JUN-2002; 2002US-00177488.

XX 26-MAR-2004; 2004US-0557116P.

XX 04-AUG-2004; 2004US-0598899P.

XX (GETH) GENENTECH INC.

XX Cairns B, Chen R, Frantz G, Hillan KJ, Koeppen H, Phillips HS;

XX Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;

XX Slukowski M;

XX WPI; 2005-384304/39.

XX P-PSDB; AEA00190.

XX Novel isolated antibody capable of binding to tumor-associated antigenic
 PT target polypeptide, useful for treating cell proliferative disorder e.g.
 PT cancer.
 XX
 XX Claim 1; SEQ ID NO 64; 337pp; English.

XX The invention relates to a novel isolated antibody binding to a
 CC polypeptide having at least 80% sequence identity to a polypeptide having
 CC any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
 CC target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
 CC specification, a polypeptide having any one of (AEA00127-AEA00202),
 CC lacking its associated signal peptide, or an extracellular domain of a
 CC polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
 CC encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
 CC An antibody of the invention has cytostatic activity. The antibody is
 CC useful for inhibiting growth of a cell expressing TAT188, which involves
 CC contacting the cell with the antibody. The cell is a cancer cell chosen
 CC from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
 CC liver cell. The cancer cell is a mammalian cell, preferably a human cell.
 CC The antibody is also useful for detecting the level of TAT188
 CC polypeptide expressed in a test cell relative to a control cell, and for
 CC detecting the level of TAT188 polypeptide or a polypeptide having at
 CC least 80 % sequence identity to the TAT188 polypeptide sequence in a test
 CC cell relative to a control cell. The antibody is useful for treating a
 CC preventing a cell proliferative disorder associated with increased
 CC expression or activity of a polypeptide having at least 80 % identity to
 CC a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
 CC The method of the invention is useful for inhibiting the growth of a
 CC cancer cell. The present sequence encodes a polypeptide of the invention.

XX SQ Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 14; Length 1460;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19

Db 1186 CTTTACTTCATAGTCTTTG 1168

RESULT 16

AEA00632/c

ID AEA00632 standard; cDNA; 1460 BP.

XX AEA00632;

AC AEA00632;

DT 28-JUL-2005 (first entry)

XX Human TAT143 cDNA sequence SeqID64.

DE antibody identification; tumor-associated antigen; cytostatic;

XX RNA interference; gene therapy; cell death; cancer; breast tumor;

XX colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;

XX skin tumor; liver tumor; gene; ss; TAT143.

XX Homo sapiens.

OS US2005107595-A1.

XX 19-MAY-2005.

XX 10-SEP-2004; 2004US-00938061.

XX 20-JUN-2001; 2001US-0299500P.

XX 29-JUN-2001; 2001US-0301880P.

XX 18-SEP-2001; 2001US-0323268P.

XX 19-JUN-2002; 2002US-00177488.

XX 26-MAR-2004; 2004US-0557116P.

XX 04-AUG-2004; 2004US-0598899P.

XX (GETH) GENENTECH INC.

XX Cairns B, Chen R, Frantz G, Hillan KJ, Koeppen H, Phillips HS;

XX Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;

XX Sakanaka C, Chuntharapai A, Reed CJ;

XX WPI; 2005-371577/38.

XX

DR P-PSDB; AEA00710.

XX Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds

PT to tumor-associated antigenic target polypeptide, useful for diagnosing

PT or treating cancer.

XX

PS Claim 1; SEQ ID NO 64; 96pp; English.

XX

XX This invention relates to a novel isolated antibody, for example anti-E16

CC or anti-TAT112 antibody, that binds to a tumor-associated antigenic

CC target polypeptide (TAT) and that lacks an associated signal peptide

CC sequence. The invention may be useful for the development of compounds

CC with a cytostatic activity acting as antagonists of the TAT188

CC polypeptide or RNA interference whilst the disclosed sequences may be

CC useful for gene therapy. The invention is useful for inducing the death

CC of a cell (such as a cancer cell chosen from breast, colon, rectum,

CC endometrium, kidney, lung, ovary, skin and liver) to which it binds,

CC inhibiting proliferation or promoting cell death of a cell expressing

CC TAT188. In addition, the invention may be useful for detecting the level

CC of TAT188 polypeptide in a test cell relative to a control cell, or

CC treating or preventing a cell proliferative disorder associated with

CC increased expression of TAT188. The novel antibody of the invention is

CC useful for inhibiting the growth of a cancer cell and may be useful for

CC diagnosing or treating cancer. The present sequence is that of the human

CC TAT143 cDNA which encodes a protein against which an antibody of the

CC invention may be targeted.

XX

SQ Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 14; Length 1460;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTTCATAGTCTTTG 19

Db 1186 CTTTACTTTCATAGTCTTTG 1168

RESULT 17

AAF58615/c

ID AAF58615 standard; cDNA; 1619 BP.

XX

AC AAF58615;

XX

DT 24-APR-2001 (first entry)

XX

DE Human RECAP polynucleotide, SEQ ID NO: 43.

XX

XX Human, RECAP; receptors and associated proteins; cerebroprotective;

KW neotropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;

KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;

KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;

KW cystostatic; antibacterial; virucide; fungicide; protozoacide;

KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200107612-A2.

XX

PD 01-FEB-2001.

XX

XX 21-JUL-2000; 2000WO-US020035.

XX

XX 21-JUL-1999; 99US-0145232P.

PR 07-OCT-1999; 99US-0158578P.

PR 12-NOV-1999; 99US-0165192P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

XX Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;

PI Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;

XX

XX WPI; 2001-168554/17.

DR P-PSDB; AAB68891.

XX

PT Novel receptors and associated proteins for diagnosis and treatment of

PT neurological disorders, immunological disorders including autoimmune/

PT inflammatory disorders and cell proliferative disorders such as cancer.

XX

PS Example 5; Page 126-127; 128pp; English.

XX

XX The present sequence encodes a human RECAP (receptors and associated

CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful

CC in the diagnosis, treatment and prevention of neurological disorders such

CC as stroke, Alzheimer's disease, Pick's disease, Huntington's disease,

CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral

CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD

CC (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker

CC syndrome); immunological disorders, including autoimmune/inflammatory

CC disorders such as AIDS, DiGeorge's syndrome, severe combined

CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's

CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,

CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,

CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,

CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and

CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,

CC cirrhosis, hepatitis and cancer

XX

SQ Sequence 1619 BP; 331 A; 463 C; 400 G; 425 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 4; Length 1619;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTTCATAGTCTTTG 19

Db 1178 CTTTACTTTCATAGTCTTTG 1160

RESULT 18

AAZ90056/c

ID AAZ90056 standard; cDNA; 1718 BP.

XX

AC AAZ90056;

XX

DT 09-MAY-2000 (first entry)

XX

DE Hydrophobic domain containing protein clone HP10549 nucleotide sequence.

XX

XX Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;

KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;

KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;

KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;

XX anti-inflammatory; infection; bodily characteristic; ss.

XX

OS Homo sapiens.

XX

PN WO200000506-A2.

XX

PD 06-JAN-2000.

XX

XX 18-JUN-1999; 99WO-JP003242.

XX

XX 26-JUN-1998; 98JP-00180008.

PR

XX (SAGA) SAGAMI CHEM RES CENT.

PA (PROT-) PROTEGENE INC.

XX

PI Kato S, Kimura T;

XX

XX WPI; 2000-160665/14.

DR

DR P-PSDB; AAY78809.

XX

XX Novel human proteins having hydrophobic domains used for research and

PT diagnostic purposes.

XX

PS Claim 4; Page 111-113; 117pp; English.

XX This sequence represents the hydrophobic domain containing protein, clone
 CC HP10549 nucleotide sequence. The sequence is isolated from a human
 CC stomach cancer cell line. The invention relates to human proteins with
 CC hydrophobic domains, the DNA and the cDNA encoding them. The
 CC polynucleotides and proteins are predicted to have biological activities
 CC which make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals. Suggested activities include
 CC nutritional activity (nutritional source or supplement); cytokine and
 CC cell proliferation/differentiation activity; immune stimulating (e.g. as
 CC vaccines) or suppressing activity (e.g. to treat various immune
 CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic
 CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation, Guillain-Barre syndrome, myasthenia gravis, graft-versus-host disease
 CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
 CC organ transplantation); haemopoiesis regulating activity (e.g. in
 CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
 CC activity (e.g. wound healing and tissue repair, ulcers, burns,
 CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
 CC activity; haemostatic and thrombolytic activity (e.g. treating
 CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
 CC tumour inhibition activity. The polynucleotides are also stated to be
 CC useful for gene therapy. Other activities include inhibiting infections
 CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
 CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
 CC skin, effecting biorhythms or circadian cycles; enhancing fertility;
 CC treatment of depression; treatment of pain; hormonal or endocrine
 CC activity. The polynucleotides may also be used for recombinant expression
 CC of the protein

XX SQ Sequence 1718 BP; 381 A; 495 C; 377 G; 465 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 3; Length 1718;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
 |||||
 Db 602 CTTTACTTCATAGTCTTTG 584

RESULT 19

ABT10173/c

ID ABT10173 standard; cDNA; 2302 BP.

XX AC ABT10173;

XX DT 04-DEC-2002 (first entry)

XX DE Human breast cancer associated coding sequence SEQ ID NO: 307.

XX KW Human; breast specific gene; breast cancer; differential expression;
 KW cytosstatic; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200259271-A2.

XX PD 01-AUG-2002.

XX PF 25-JAN-2002; 2002WO-US002176.

XX PR 25-JAN-2001; 2001US-0263757P.

XX PR 25-APR-2001; 2001US-0286090P.

XX PR 23-MAY-2001; 2001US-0292517P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Orr MS, Nation M, Diggins JC, Zeng W;

XX WP1; 2002-674803/72.

XX

PT Diagnosing breast cancer in a patient comprises detecting the level of
 PT gene expression in cell or tissue samples, where a differential gene
 PT expression is indicative of breast cancer.

XX PS Claim 1; SEQ ID NO 307; 260pp + Sequence Listing; English.

XX The present invention relates to methods of diagnosing breast cancer in a
 CC patient, which comprise detecting the level of expression in a tissue
 CC sample of two or more genes selected from those shown in ABT09867-
 CC ABT1112, where a differential expression of the genes indicates breast
 CC cancer. The methods are useful in diagnosing, treating, detecting the
 CC progression, and in monitoring treatment of breast cancer in patients.
 CC The methods are also useful as a screening tool for agents that modulate
 CC the onset or progression of breast cancer. The breast cancer genes may be
 CC used as diagnostic markers for the prediction or identification of the
 CC malignant state of breast tissue, for confirming the type and progression
 CC of cancer, and for drug screening and assays. The present sequence is a
 CC coding sequence of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub.published_pct_sequences

XX SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 2302;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
 |||||
 Db 1158 CTTTACTTCATAGTCTTTG 1140

RESULT 20

ACC58385/c

ID ACC58386 standard; cDNA; 2302 BP.

XX AC ACC58386;

XX DT 26-AUG-2003 (first entry)

XX DE Human GPCR-like retinoic acid-induced gene 1.

XX KW Human; retinoic acid-induced gene 1; RAIG1; feeding; fasting; GPCR;
 KW receptor; G-protein coupled receptor; anorectic; antidiabetic;
 KW antidepressant; immunomodulator; transgenic; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX CDS 100..1173

XX FT /*tag= a

XX FT /product= "GPCR-like RAIG1"

XX PN WO2003016553-A2.

XX PD 27-FEB-2003.

XX PF 20-AUG-2002; 2002WO-US026510.

XX PR 20-AUG-2001; 2001US-0313940P.

XX PA (GETH) GENENTECH INC.

XX PA (CURA-) CURAGEN CORP.

XX PI Lewin DA, Stewart TA;

XX WP1; 2003-278580/27.

XX P-PSDB; ABR42649.

XX New G-protein coupled receptor-like retinoic acid induced gene 1 (GPCR-
 PT like RAIG1) polypeptide and gene, useful for diagnosing or treating

PT metabolic disorders, e.g. obesity, anorexia, cachexia or diabetes.
XX Claim 10; Page 16-17; 150pp; English.
PS
XX
CC The present sequence is that of human G-protein coupled receptor-like
CC retinoic acid induced gene 1 (GPCR-like RAIG1). This is the human
CC homologue of murine GPCR-like RAIG1. The murine gene was shown to be
CC differentially regulated during fasting-feeding cycles in mice, with
CC moderate induction early in fasting, down-regulation with extended
CC fasting and 4-fold up-regulation with feeding in recovery from fasting.
CC The differentially expressed gene, its mRNA, and the encoded protein, can
CC each be manipulated to detect and treat metabolic disorders associated
CC with up- or down-regulation of GPCR-like RAIG1 activity, such as obesity,
CC anorexia, cachexia or diabetes
XX
SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 8; Length 2302;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 1158 CTTTACTTCATAGTCTTTG 1140
RESULT 21
ADD93240/c
ID ADD93240 standard; cDNA; 2302 BP.
AC ADD93240;
XX
DT 29-JAN-2004 (first entry)
XX
DE RAIG1 coding sequence.
XX
XX ss; gene; retinoic acid-inducible gene 1 ; RAIG1; orphan ;
KW G-protein coupled receptor; GPCR; chromosome 12; GPCR5B; carcinoma ;
KW vaccine; breast cancer; pancreatic cancer; lung cancer; liver cancer;
KW ovarian cancer; colon cancer ; osteosarcoma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 100..1173
FT /*tag= a
FT /*product= "RAIG1"
XX
XX WO2003087832-A2.
XX
XX 23-OCT-2003.
XX
XX 10-APR-2003; 2003WO-GB001587.
XX
XX 11-APR-2002; 2002GB-00008331.
XX
XX 17-SEP-2002; 2002GB-00021538.
XX
XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.
XX
XX Terrett JA;
XX
XX WPI; 2003-845382/78.
XX
XX P-PSDB; ADD93239.
XX
XX GENBANK; AF095448.
XX
XX Screening, diagnosing and/or treating carcinoma, including breast,
XX pancreatic, lung, liver, ovarian and colon cancer by detecting the change
XX in expression or activity of an RAIG1 polypeptide or encoding nucleic
XX acid molecule.
XX
XX Claim 1; Fig 2; 43pp; English.
XX
XX This sequence encodes retinoic acid-inducible gene 1 (RAIG1) polypeptide.

CC RAIG1 is an orphan G-protein coupled receptor (GPCR) located on
CC chromosome 12. RAIG1 shows a restricted expression pattern compared to a
CC related receptor, GPCR5B, which is widely expressed in peripheral and
CC central tissues. The RAIG1 cDNA and polypeptide sequence may be used in
CC the method of the invention for screening for and/or diagnosis of
CC carcinoma in a subject, and/or monitoring the effectiveness of carcinoma
CC therapy. The method comprises detecting and/or quantifying in a
CC biological sample obtained from the subject an RAIG1 polypeptide and a
CC nucleic acid molecule. The RAIG1 polypeptide and nucleic acid molecule
CC are useful in the manufacture of a medicament for the treatment of
CC carcinoma, where the composition is a vaccine. An agent which interacts
CC with or causes change in the expression or activity of an RAIG1
CC polypeptide or nucleic acid molecule, is also useful in the manufacture
CC of a medicament for the treatment of carcinoma that is breast cancer,
CC pancreatic cancer, lung cancer, liver cancer, ovarian cancer, colon
CC cancer and/or osteosarcoma. They can also be used in the diagnosis and
CC screening of such carcinomas.
XX
SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 1158 CTTTACTTCATAGTCTTTG 1140
RESULT 22
ADL24773/c
ID ADL24773 standard; DNA; 2302 BP.
XX
AC ADL24773;
XX
DT 20-MAY-2004 (first entry)
XX
DE Intestinal epithelium/peyer's patch M cell-associated DNA sequence #171.
XX
XX intestinal epithelium cell development; peyer's patch M cell development;
XX inflammatory bowel disease; glutenenteropathy; infectious disease;
KW autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;
KW Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;
KW immune system disorder; hypersensitivity; anaphylaxis;
KW blood group incompatibility; ds; human.
XX
XX Homo sapiens.
XX
XX WO200280852-A2.
XX
XX 17-OCT-2002.
XX
XX 04-APR-2002; 2002WO-US010873.
XX
XX 04-APR-2001; 2001US-0281416P.
XX
XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
XX Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;
XX
XX WPI; 2003-075470/07.
XX
XX Novel isolated or purified polypeptide encoded by genes associated with
XX intestinal epithelium or M cell development, differentiation or function,
XX useful for treating autoimmune diseases and infectious diseases.
XX
XX Claim 1; SEQ ID NO 283; 152pp; English.
XX
XX The invention comprises DNA sequences which are associated with
XX intestinal epithelium and peyer's patch M cells. The DNA sequences of the
XX invention are useful for assessing, modifying, modulating or regulating
XX intestinal epithelium or M cell development. The DNA sequences of the
XX invention are also useful in the treatment of: inflammatory bowel

CC disease, glutenenteropathy, infectious diseases, autoimmune diseases
 CC (e.g. haemolytic anemia, rheumatoid arthritis, dermatitis, Grave's
 CC diseases, multiple sclerosis, allergy, asthma and diabetic mellitus),
 CC diseases or disorders of the immune system, hypersensitivity,
 CC anaphylaxis, and blood group incompatibility. The present nucleic acid
 CC represents an intestinal epithelium/peyer's patch M cell-associated DNA
 CC sequence of the invention. NOTE: The present sequence is not shown in the
 CC specification, but has been retrieved from the WIPO website.
 XX
 SQ Sequence 2302 BP; 494 A; 566 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 2302;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTG 19
 Db 1158 CTTTACTTCATAGCTCTTG 1140

RESULT 23

ABT31923/C

ID ABT31923 standard; DNA; 2316 BP.

XX AC ABT31923;

XX AC (first entry)

XX 01-MAY-2003 Human breast cancer / ovarian cancer related coding sequence #30.

DE Human; gene; ds; cytostatic; breast cancer; ovarian cancer.

XX Homo sapiens.

XX WO2003000012-A2.

XX 03-JAN-2003.

XX 21-JUN-2002; 2002WO-US019773.

XX 21-JUN-2001; 2001US-0300159P.

PR 27-JUN-2001; 2001US-0301351P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Vaiby OP;

XX WPI; 2003-267848/26.

DR P-PSDB; ABJ37054.

XX Determining the presence of breast cancer in an individual, involves
 PT using specific polynucleotide markers.

XX Disclosure; Page 163; 233pp; English.

XX The invention comprises a method for assessing whether a patient is
 CC afflicted with breast cancer or ovarian cancer. The method involves the
 CC use of specific DNA markers. The method of the invention is useful in the
 CC detection and treatment of ovarian and breast cancer. DNA sequences
 CC ABT31894 - ABT31949 encode human breast/ovarian cancer-related proteins
 XX

SQ Sequence 2316 BP; 502 A; 670 C; 535 G; 609 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 2316;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTG 19
 Db 1164 CTTTACTTCATAGCTCTTG 1146

RESULT 24

AAH14688/C
 ID AAH14688 standard; cDNA; 2446 BP.
 AC AAH14688;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human cDNA sequence SEQ ID NO:12388.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

PS Claim 8; SEQ ID NO 12388; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX Sequence 2446 BP; 507 A; 723 C; 572 G; 644 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 4; Length 2446;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTG 19
 Db 1312 CTTTACTTCATAGCTCTTG 1294

RESULT 25
AB242832/c
ID AB242832 standard; DNA; 2456 BP.
XX AC AB242832;
XX DT 04-MAR-2003 (first entry)
XX DE Human G protein-coupled receptor RAIG1 nucleotide SEQ ID NO:453.
XX DE
XX KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
XX KW G protein-coupled receptor modulator; antibody; immune-related disease;
XX KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
XX KW immunological-related cell proliferative disease; autoimmune disease;
XX KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
XX KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
XX KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
XX KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
XX KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
XX KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
XX KW ulcer; gene; ds.
XX OS Homo sapiens.
XX PN WO200261087-A2.
XX PD 08-AUG-2002.
XX PF 19-DEC-2001; 2001WO-US050107.
XX PR 19-DEC-2000; 2000US-0257144P.
XX PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX PI Burmer GC, Roush CL, Brown JP;
XX WP1; 2003-046718/04.
XX P-PSDB; ABP81984.
XX New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.
XX Disclosure; Fig 1; 523pp; English.
XX The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83819) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention
XX SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
DB 1312 CTTTACTTCATAGTCTTTG 1294
RESULT 26
ACC72695/c
ID ACC72695 standard; cDNA; 2456 BP.
XX AC ACC72695;
XX DT 09-JUL-2003 (first entry)
XX DE Human cancer related protein encoding cDNA SEQ ID NO:34.
XX KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
XX KW heart disease; atherosclerosis; endometriosis; gene; ss.
XX OS Homo sapiens.
XX PN WO2003025138-A2.
XX PD 27-MAR-2003.
XX PF 17-SEP-2002; 2002WO-US029560.
XX PR 17-SEP-2001; 2001US-0323469P.
XX PR 20-SEP-2001; 2001US-0323887P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 08-FEB-2002; 2002US-0355145P.
XX PR 08-FEB-2002; 2002US-0355257P.
XX PR 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX AFar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
XX Zlotnik A;
XX WPI; 2003-354600/33.
XX P-PSDB; ABR58568.
XX New genes that are up-regulated or down-regulated in cancers, useful as markers for diagnosing e.g. cancer, ischemia or heart diseases, or as therapeutic targets for screening drugs for treating these diseases.
XX Claim 8; Page 643-644; 767pp; English.
XX The present invention describes an isolated nucleic acid molecule, which comprises the sequence of any of the genes that are up-regulated or down-regulated in specific cancers (e.g. about 1031 genes up-regulated in acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer related gene nucleotide sequences which encode the proteins given in ABR58521 to ABR58709. Also described: (1) determining the presence or absence of a pathological cell in a patient; (2) an expression vector comprising a nucleic acid molecule described above; (3) a host cell comprising the vector; (4) an isolated polypeptide, which is encoded by the nucleic acid; (5) an antibody that specifically binds the polypeptide of (4); (6) specifically targeting a compound to a pathological cell in a patient by administering to the patient the antibody above; and (7) a drug screening assay. The nucleic acid is useful as diagnostic markers or therapeutic targets. In particular, the nucleic acid is useful for diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow, bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary, pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases, atherosclerosis and endometriosis. The nucleic acid is also useful in drug screening, particularly for identifying agents for treating these pathologies

SQ Sequence 2456 BP; 522 A; 720 C; 571 G; 643 T; 0 U; 0 Other;
 Query Match 100.0%; Score 19; DB 10; Length 2456;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGCTTTG 19
 Db 1312 CTTTACTTCATAGCTTTG 1294
 RESULT 27
 ADN39301/c
 ID ADN39301 standard; cDNA; 2456 BP.
 XX AC ADN39301;
 XX DT 17-JUN-2004 (first entry)
 XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:619.
 XX KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnary; gene therapy; vaccine; gene; ss.
 XX OS Homo sapiens.
 XX PN W02003042661-A2.
 XX PD 22-MAY-2003.
 XX PF 13-NOV-2002; 2002WO-US036810.
 XX PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-035250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-036809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX WIPI; 2003-468649/44.
 DR P-PSDB; ADN39302.
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX Claim 8; SEQ ID NO 619; 1385pp; English.
 XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
 Query Match 100.0%; Score 19; DB 11; Length 2456;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGCTTTG 19
 Db 1312 CTTTACTTCATAGCTTTG 1294
 RESULT 28
 ADI28525/c
 ID ADI28525 standard; cDNA; 2456 BP.
 XX AC ADI28525;
 XX DT 22-APR-2004 (first entry)
 XX DE Human GPCR retinoic acid induced 3 (RAI-3) cDNA.
 XX KW Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
 KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
 KW broncholytic; gene therapy; gene; single nucleotide polymorphism; SNP;
 KW chromosome 12p13-p12.3; ss.
 XX OS Homo sapiens.
 PH Key Location/Qualifiers
 FT variation /tag= a
 FT /label= RAI-3-s1
 FT /note= "located in 5' untranslated region"
 FT /standard_name= "Single nucleotide polymorphism"
 FT CDS 254..1327
 FT /tag= b
 FT /product= "Human RAI3"
 FT variation replace(364,y)
 FT /tag= c
 FT /label= RAI-3-s2
 FT /note= "exon 1, silent (Ala/Ala)"
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(511,y)
 FT /tag= d
 FT /label= RAI-3-s3
 FT /note= "exon 2, silent (Ile/Ile)"
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(523,y)
 FT /tag= e
 FT /label= RAI-3-s4
 FT /note= "exon 2, silent (Asp/Asp)"
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(605,r)
 FT /tag= f
 FT /label= RAI-3-s6
 FT /note= "exon 2, missense (Ser/Gly)"
 FT /standard_name= "Single nucleotide polymorphism"
 FT variation replace(797,r)

Fi /note= "missense (Gln/Arg)"
FT /standard_name= "single nucleotide polymorphism"
XX
XX WO2004001060-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019255.
XX
XX 20-JUN-2002; 2002US-0390850P.
XX
XX 29-AUG-2002; 2002US-0407006P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Whitney GS, Opitck G, Garulacan L, Ramanathan CS, McKinnon M;
XX Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
XX
XX WPI; 2004-090973/09.
XX P-PSDB; ADI28460.
XX GENBANK; NM_003979.
XX
XX New nucleic acid molecule encoding a human G-protein coupled receptor,
XX RAI3, useful for preventing, treating or ameliorating chronic obstructive
XX pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
XX of COPD.
XX
XX Claim 1; SEQ ID NO 2; 301pp; English.
XX
XX The present sequence is that of cDNA encoding a human G-protein coupled
XX receptor (Class C, Group 5), denoted retinoic acid induced 3 (RAI-3).
XX Proteomic methods were used to isolate cigarette smoke-inducible
XX tyrosine phosphorylated proteins from airway epithelial cells. RAI-3 was
XX identified as being tyrosine phosphorylated and/or as being
XX associated/complexed with tyrosine phosphorylated proteins only in those
XX cells that had been exposed to cigarette smoke. Since RAI-3 is primarily
XX expressed in lung tissue, and since cigarette smoke is a major causative
XX factor of chronic obstructive pulmonary disease (COPD), RAI-3 provides a
XX novel cellular target for identifying modulators, e.g. agonists or
XX antagonists, useful for the treatment and/or prevention of COPD and
XX related disorders such as emphysema and chronic bronchitis. RAI-3
XX modulators, e.g. agonists and antagonists, especially antisense
XX compounds, can be used to treat COPD and other disorders and diseases
XX associated with regulation of NF-kB and/or its associated or interacting
XX signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
XX the RAI-3 gene are useful for determining COPD association in
XX individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
XX preventing, treating or ameliorating disorders related to aberrant GPCR
XX signaling or cell cycle regulation, pulmonary disorders, inflammatory
XX lung disorders, COPD, the underlying symptoms of COPD, COPD-related
XX disorders or conditions, autoimmune disorders, disorders related to
XX hyperimmune activity, inflammatory conditions, disorders related to
XX aberrant acute phase responses, hypercongenital conditions, birth
XX defects, necrotic lesions, wounds, organ transplant rejection, renal
XX diseases, ischemia-reperfusion injury, heart disorders, disorders
XX related to aberrant signal transduction, proliferation disorders,
XX cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
XX ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
XX nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
XX rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
XX pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
XX colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
XX associated with aberrant cell adhesion, I-CAM function and/or regulation,
XX E-selectin function and/or regulation, or aberrant NF-kB function and/or
XX regulation (all claimed).
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 19; DB 12; Length 2456;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
|||||

Db 1312 CTTTACTTCATAGTCTTTG 1294
RESULT 30
ADR48221/C
ID ADR48221 standard; cDNA; 2456 BP.
XX
XX ADR48221;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human retinoic acid induced 3 (RAI3) encoding cDNA SEQ:9.
XX
XX pancreatic cancer-associated transcript; pancreatic cancer; human;
XX cytosatic; gene therapy; protein therapy; retinoic acid induced 3; RAI3;
XX gene; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX CDS 254..1327
XX FT /*tag= a
XX FT /product= "retinoic acid induced 3 (RAI3)"
XX
XX WO2004074510-A1.
XX
XX 02-SEP-2004.
XX
XX 18-FEB-2004; 2004WO-AU000194.
XX
XX 18-FEB-2003; 2003AU-00900747.
XX
XX (GARV-) GARVAN INST MEDICAL RES.
XX
XX Biankin A, Segara D, Henshall S, Sutherland R;
XX
XX WPI; 2004-635591/61.
XX P-PSDB; ADR48222.
XX
XX Detecting pancreatic cancer-associated transcript in a biological sample,
XX useful for diagnosing or treating the disease, comprises contacting the
XX sample with a polynucleotide that selectively hybridizes to a specific
XX sequence.
XX
XX Claim 7; SEQ ID NO 9; 263pp; English.
XX
XX The present invention describes a method for detecting a pancreatic
XX cancer-associated transcript in a biological sample. The method comprises
XX contacting the biological sample with a polynucleotide that selectively
XX hybridizes to a sequence at least 80% identical to a sequence as shown in
XX any one of Tables 3 to 25 in the specification or having the GenBank
XX Accession Number AF279145. Also described: (1) diagnosing pancreatic
XX cancer in a human or animal subject being tested, determining the
XX likelihood that a subject having a pancreatic cancer will survive, or
XX determining the suitability of a subject having a pancreatic cancer for
XX surgical resection therapy; (2) detecting a pancreatic cancer-associated
XX polypeptide in a biological sample; (3) determining the likelihood that a
XX subject having a pancreatic cancer will survive; and (4) monitoring the
XX efficacy of a therapeutic treatment of pancreatic cancer. A pancreatic
XX cancer-associated transcript has cytostatic activity, and can be used in
XX gene and protein therapy. A pancreatic cancer-associated transcript
XX polynucleotide, a vector comprising the polynucleotide, an isolated
XX polypeptide or an antibody that binds to the isolated polypeptide can be
XX used for diagnosing or prognosing pancreatic cancer or for preparing a
XX medicament for the treatment of pancreatic cancer. The prognostic or
XX diagnostic methods are useful for the early detection of pancreatic
XX cancer or its metastases, and for monitoring the progress of disease
XX as during remission or following surgery or chemotherapy. The present
XX sequence encodes human retinoic acid induced 3 (RAI3), which is used in
XX the exemplification of the present invention.
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
SQ

Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||
1312 CTTTACTTCATAGTCTTTG 12

RESULT 33
ADU06126/C
ID ADU06126 standard; DNA; 2456 BP.
XX
XX ADU06126;
AC
XX
XX 27-JAN-2005 (first entry)
DT
XX
XX Novel bronchial cancer-associated human gene SeqID350.
DE
XX
XX bronchial cancer; cytostatic; tumour-associated protein;
KW cancer detection; metastasis; tumour; gene; ds; human.
KW
XX
XX Homo sapiens.
OS
XX
XX DE10316701-A1.
FN
XX
XX 04-NOV-2004.
PD
XX
XX 09-APR-2003; 2003DE-01016701.
PF
XX
XX 09-APR-2003; 2003DE-01016701.
PR
XX
XX (HINZ/) HINZMANN B.
PA (HERM/) HERMANN K.
PA (CAST/) HEIDEN CASTANOS-VELEZ E.
XX
XX Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pillarsky C;
PI
XX
XX WPI: 2004-786403/78.
DR P-PSDB; ADU06613.
XX
XX New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT cancer and in screening for therapeutic and diagnostic agents.
PT
XX
XX Claim 1; SEQ ID NO 350; 1381bp; German.
PS
XX
XX This invention relates to a novel isolated nucleic acid associated with
CC bronchial cancer comprising 489 defined sequences given in the
CC specification. The invention may be useful for the production of
CC compounds with a cytostatic activity through the inhibition of expression
CC or activity of tumour-associated proteins. The novel DNA sequences and
CC the proteins/peptides encoded by them are used for detecting bronchial
CC cancer or determining the risk of developing it and to screen for
CC specific binding partners of the DNA or protein sequences, where the
CC binding partners are potentially useful as agents for treating or
CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a novel
CC bronchial cancer-associated human gene sequence of the invention.
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 13; Length 2456;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGCTTTG 19
|||||
Db 1312 CTTTACTTCATAGCTTTG 1294
|||||
RESULT 34
AEA00088/C
ID AEA00088 standard; cDNA; 2456 BP.
XX
XX AEA00088;
AC
XX
XX 28-JUL-2005 (first entry)
DT

XX
DE
XX
KW ss; gene; tumor-associated antigen; cytostatic; breast tumor;
KW endocrine disease; gynecology and obstetrics; neoplasm; colon tumor;
KW gastrointestinal disease; rectal tumor; endometroid carcinoma;
KW genitourinary disease; renal tumor; lung tumor; respiratory disease;
KW ovary tumor; skin tumor; liver tumor.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 254..1327
FT /*tag= a
FT /product= "TAT115"
XX
XX US2005106644-A1.
FN
XX
XX 19-MAY-2005.
PD
XX
XX 08-SEP-2004; 2004US-00936626.
PF
XX
XX 20-JUN-2001; 2001US-0299500P.
PR 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323268P.
PR 19-JUN-2002; 2002US-00177488.
PR 26-MAR-2004; 2004US-0557116P.
PR 04-AUG-2004; 2004US-0598899P.
XX
XX (GETH) GENENTECH INC.
PA
XX
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koeppen H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sliwkowski M;
PI
XX
XX WPI: 2005-384304/39.
DR P-PSDB; AEA00166.
XX
XX Novel isolated antibody capable of binding to tumor-associated antigenic
PT target polypeptide, useful for treating cell proliferative disorder e.g.
PT cancer.
XX
XX Claim 1; SEQ ID NO 40; 337pp; English.
PS
XX
XX The invention relates to a novel isolated antibody binding to a
CC polypeptide having at least 80% sequence identity to a polypeptide having
CC any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
CC target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
CC specification, a polypeptide having any one of (AEA00127-AEA00202),
CC lacking its associated signal peptide, or an extracellular domain of a
CC polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
CC encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
CC An antibody of the invention has cytostatic activity. The antibody is
CC useful for inhibiting growth of a cell expressing TAT188, which involves
CC contacting the cell with the antibody. The cell is a cancer cell chosen
CC from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
CC liver cell. The cancer cell is a mammalian cell, preferably a human cell.
CC The antibody is also useful for detecting the level of TAT188
CC polypeptide expressed in a test cell relative to a control cell, and for
CC detecting the level of TAT188 polypeptide or a polypeptide having at
CC least 80 % sequence identity to the TAT188 polypeptide sequence in a test
CC cell relative to a control cell. The antibody is useful for treating a
CC preventing a cell proliferative disorder associated with increased
CC expression or activity of a polypeptide having at least 80 % identity to
CC a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
CC The method of the invention is useful for inhibiting the growth of a
CC cancer cell. The present sequence encodes a polypeptide of the invention.
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
Query Match 100.0%; Score 19; DB 14; Length 2456;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 CTTTACTTCATAGCTTTG 19
      |||||
Db     1312 CTTTACTTCATAGCTTTG 1294

RESULT 35
AEA00608/c
ID AEA00608 standard; cDNA; 2456 BP.
XX
AC AEA00608;
XX
DT 28-JUL-2005 (first entry)
XX
DE Human TAT115 cDNA sequence SeqID40.
XX
KW antibody identification; tumor-associated antigen; cytostatic;
KW RNA interference; gene therapy; cell death; cancer; breast tumor;
KW colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;
KW skin tumor; liver tumor; gene; ss; TAT115.
XX
OS Homo sapiens.
XX
PN US2005107595-A1.
XX
PD 19-MAY-2005.
XX
PF 10-SEP-2004; 2004US-00938061.
XX
PR 20-JUN-2001; 2001US-0299500P.
PR 29-JUN-2001; 2001US-0301880P.
PR 18-SEP-2001; 2001US-0323268P.
PR 19-JUN-2002; 2002US-00177488.
PR 26-MAR-2004; 2004US-0557116P.
PR 04-AUG-2004; 2004US-0598899P.
XX
PA (GETH ) GENENTECH INC.
XX
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koeppe H, Phillips HS;
PI Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
PI Sakanaka C, Chuntharapai A, Reed CJ;
XX
DR WPI; 2005-371577/38.
DR P-PSDB; AEA00686.
XX
XX Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds
PT to tumor-associated antigenic target polypeptide, useful for diagnosing
PT or treating cancer.
XX
XX Claim 1; SEQ ID NO 40; 96pp; English.
XX
XX This invention relates to a novel isolated antibody, for example anti-E16
CC or anti-TAT112 antibody, that binds to a tumor-associated antigenic
CC target polypeptide (TAT) and that lacks an associated signal peptide
CC sequence. The invention may be useful for the development of compounds
CC with a cytostatic activity acting as antagonists of the TAT118
CC polypeptide or RNA interference whilst the disclosed sequences may be
CC useful for gene therapy. The invention is useful for inducing the death
CC of a cell (such as a cancer cell chosen from breast, colon, rectum,
CC endometrium, kidney, lung, ovary, skin and liver) to which it binds,
CC inhibiting proliferation or promoting cell death of a cell expressing
CC TAT118. In addition, the invention may be useful for detecting the level
CC of TAT118 polypeptide in a test cell relative to a control cell, or
CC treating or preventing a cell proliferative disorder associated with
CC increased expression of TAT118. The novel antibody of the invention is
CC useful for inhibiting the growth of a cancer cell and may be useful for
CC diagnosing or treating cancer. The present sequence is that of the human
CC TAT115 cDNA which encodes a protein against which an antibody of the
CC invention may be targeted.
XX
SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
      Query Match      100.0%; Score 19; DB 14; Length 2456;
```

Best Local Similarity 100.0%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTTTG 19
 |||||
Db 1312 CTTTACTTCATAGCTTTG 1294

RESULT 36
AED47480/c
ID AED47480 standard; DNA; 2456 BP.
XX
AC AED47480;
XX
DT 15-DEC-2005 (first entry)
XX
DE Retinoic acid-inducible G-protein coupled receptor 3 gene, SEQ ID 2.
XX
KW RNA Interference; gene silencing;
KW retinoic acid-inducible G-protein coupled receptor 3; breast tumor;
KW cytostatic; cell growth; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 254..1327
FT /*tag= a
FT /product= "Retinoic acid-inducible G-protein coupled
FT receptor 3"
FT misc_feature 338..403
FT /*tag= b
FT /label= Transmembrane domain
FT misc_feature 452..514
FT /*tag= c
FT /label= Transmembrane domain
FT misc_feature 497..1057
FT /*tag= d
FT /label= 7 transmembrane receptor
FT misc_feature 548..601
FT /*tag= e
FT /label= Transmembrane domain
FT misc_feature 656..718
FT /*tag= f
FT /label= Transmembrane domain
FT misc_feature 776..832
FT /*tag= g
FT /label= Transmembrane domain
FT misc_feature 887..946
FT /*tag= h
FT /label= Transmembrane domain
FT polyA_signal 2403..2408
FT /*tag= i
FT polyA_site 2442..2456
FT /*tag= j
XX JP2005287456-A.
XX 20-OCT-2005.
XX 02-APR-2004; 2004JP-00110232.
XX 02-APR-2004; 2004JP-00110232.
XX (NIPK) NIPPON KAYAKU KK.
XX (UVNI-) UNIV NIPPON IKA.
XX Nagahata T, Emi M;
XX WPI; 2005-737982/76.
XX P-PSDB; AED47479.
XX Substance that suppresses expression of retinoic acid-inducible G-protein
PT coupled receptor-3 RAI3 gene or nucleic acid, useful for suppressing cell

PT growth, and as pharmaceutical or therapeutic agent of breast cancer.
 XX Claim 1; SEQ ID NO 2; 22pp; Japanese.
 XX
 CC The invention relates to a novel substance which suppresses the
 CC expression of a retinoic acid-inducible G-protein coupled receptor (RAI) 3
 CC gene. The RAI3 gene suppressing agent is useful for suppressing cell
 CC growth, and as a pharmaceutical or therapeutic agent of breast cancer,
 CC and is also useful in identifying a cell growth inhibitory substance. The
 CC RAI3 gene suppressing agent enables the identification of a cell growth
 CC inhibitory substance. This polynucleotide sequence represents the
 CC retinoic acid-inducible G-protein coupled receptor 3 gene of the
 CC invention.
 XX
 SQ Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
 Query Match 100.0%; Score 19; DB 14; Length 2456;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGTCTTTG 19
 Db 1312 CTTTACTTCATAGTCTTTG 1294
 RESULT 37
 ABQ54954/c
 ID ABQ54954 standard; cDNA; 2593 BP.
 XX
 AC ABQ54954;
 XX
 XX 22-AUG-2002 (first entry)
 DT
 XX Human ovarian antigen HSKJC61 cDNA, SEQ ID NO:834.
 DE
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200200677-A1.
 XX
 XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US018569.
 XX
 XX 07-JUN-2000; 2000US-0209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX
 XX WPI; 2002-147878/19.
 XX P-PSDB; ABP41877.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 XX Claim 1; SEQ ID NO 834; 2922pp; English.
 PS
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2593 BP; 591 A; 730 C; 605 G; 667 T; 0 U; 0 Other;
 Query Match 100.0%; Score 19; DB 6; Length 2593;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGTCTTTG 19
 Db 1424 CTTTACTTCATAGTCTTTG 1406
 RESULT 38
 AAH72766/c
 ID AAH72766 standard; cDNA; 3371 BP.
 XX
 AC AAH72766;
 XX
 DT 19-SEP-2001 (first entry)
 XX
 XX Human cervical cancer marker nucleic acid 4040.
 DE
 XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 KW
 XX Homo sapiens.
 OS
 XX WO200142467-A2.
 XX
 XX 14-JUN-2001.
 XX
 XX 08-DEC-2000; 2000WO-US033312.
 XX
 XX 08-DEC-1999; 99US-0169681P.
 XX 21-DEC-1999; 99US-0171350P.
 XX 14-MAR-2000; 2000US-0189315P.
 XX 12-MAY-2000; 2000US-0203791P.
 XX 09-JUN-2000; 2000US-0210600P.
 XX 21-JUL-2000; 2000US-0220114P.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA
 XX Schlegel R, Deeds J, Berger A, Zhao X;
 XX WPI; 2001-375006/39.
 DR
 XX New isolated nucleic acid for diagnosing and treating cervical cancer and
 PT for assessing and detecting compounds for treating the cancer.

```
XX PS Claim 1; Page 840; 1051pp; English.
XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with
XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded
XX CC polypeptides are useful: to assess if a patient is afflicted with
XX CC cervical cancer or has a pre-malignant condition; to monitor the
XX CC progression of cervical cancer or a premalignant condition in a patient;
XX CC and to select and/or assess the efficacy of a compound or therapy for
XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be
XX CC useful for gene therapy
XX SQ Sequence 3371 BP; 741 A; 902 C; 779 G; 946 T; 0 U; 3 Other;

Query Match 100.0%; Score 19; DB 4; Length 3371;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
Db 1225 CTTTACTTCATAGTCCTTTG 1207

RESULT 39
ACN89274/c
ID ACN89274 standard; DNA; 4239 BP.
XX AC ACN89274;
XX DT 02-DEC-2004 (first entry)
XX DE Breast cancer related marker, seq id 10424.
XX KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX OS Homo sapiens.
XX PN US2003099974-A1.
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX PD 29-MAY-2003.
XX PF 18-JUL-2002; 2002US-00198846.
XX PR 18-JUL-2001; 2001US-0306220P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX PD WPI; 2003-787014/74.
XX PT Novel isolated polypeptide associated with breast cancer, useful for
XX PT detecting presence of polypeptide in sample, as a marker for breast
XX PT cancer.
XX PS Disclosure; SEQ ID NO 10424; 36pp; English.
XX CC The invention relates to an isolated polypeptide (I) associated with
XX CC breast cancer which is encoded by a nucleic acid molecule comprising a
XX CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX CC the polypeptide of the invention. The activity of the polypeptide of the
XX CC invention may be described as cytostatic. The antibody is useful for
XX CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX CC invention are useful in the detection of breast tumours. (I) is useful as
XX CC a marker for breast cancer and in breast cancer therapy. Sequences given
XX CC in records ACN78851-ACN92934 represent nucleic acid markers associated
XX CC with breast cancer. Note: The sequence listing does not form part of the
XX CC specification but may be obtained in electronic format from the USPTO web
XX CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX SQ Sequence 4239 BP; 859 A; 1161 C; 1075 G; 1141 T; 0 U; 3 Other;

Query Match 100.0%; Score 19; DB 11; Length 4239;

Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
Db 1720 CTTTACTTCATAGTCCTTTG 1702

RESULT 40
ADX97494/c
ID ADX97494 standard; DNA; 6730 BP.
XX AC ADX97494;
XX DT 21-APR-2005 (first entry)
XX DE Pancreatic cancer associated human protein encoding DNA, SEQ ID 42.
XX KW pancreas tumor; cytostatic; gene; ds.
XX OS Homo sapiens.
XX PN EP1471075-A2.
XX PD 27-OCT-2004.
XX PF 31-MAR-2004; 2004EP-00090124.
XX PR 31-MAR-2003; 2003DE-01015834.
XX PA (HINZ/) HINZMANN B.
XX PA (ROSE/) ROSENTHAL A.
XX PA (PILA/) PILARSKY C.
XX PA (DAHL/) DAHL E.
XX PA (SPEC/) SPECHT T.
XX PA (LICH/) LICHTNER R.
XX PI Rosenthal A, Pilarisky C, Dahl E, Specht T, Bruemendorf T;
XX PI Lichtner R, Staub E, Roepcke S, Li X;
XX PD WPI; 2004-768082/76.
XX PF P-PSDB; ADX97565.
XX PT New nucleic acid differentially expressed in pancreatic tumor tissue, for
XX PT use as diagnostic agents and in screening for therapeutic agents.
XX PS Claim 1; SEQ ID NO 42; 28pp; German.
XX CC The invention relates to a novel human nucleic acid sequence of the
XX CC pancreas and its encoded protein. The invention further comprises:
XX CC proteins and peptides, preferably isolated, that contain a sequence
XX CC encoded by the novel nucleic acid; and methods for diagnosis and
XX CC treatment of pancreatic cancer, using a substance that inhibits or binds
XX CC to the protein or its DNA, including: an antisense oligonucleotide, short
XX CC interfering RNA or ribozyme directed against the pancreatic protein, an
XX CC organic molecule, particularly having a molecular weight below 5000,
XX CC especially 300, that binds to the pancreatic DNA, an aptamer or
XX CC (monoclonal) antibody, preferably human or humanized, that binds to the
XX CC pancreatic DNA, or an anti-idiotypic antibody raised against the
XX CC monoclonal antibody, any of which may be derivatized with a reporter
XX CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
XX CC human pancreatic proteins and their encoding DNA have cytostatic
XX CC activity. The novel sequences are useful for inhibiting transcription
XX CC and/or expression of genes and proteins associated with pancreatic
XX CC cancer. This polynucleotide sequence represents the DNA encoding one of
XX CC the novel human pancreatic proteins of the invention. Note: This sequence
XX CC is not shown in the specification, it has been electronically downloaded
XX CC from a DVD-rom provided with this specification by the European Patent
XX CC Office.
XX SQ Sequence 6730 BP; 1469 A; 1763 C; 1555 G; 1943 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 13; Length 6730;
```

* Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
Db 1312 CTTTACTTCATAGCTTTG 1294

RESULT 41
AAA96701
ID AAA96701 standard; DNA; 1725 BP.
XX AC AAA96701;
XX KW
XX KW
DT 19-FEB-2001 (first entry)
XX DE
DE Reporter gene YGR131W from an ergosterol pathway.
XX Reporter gene; ergosterol pathway; protein kinase C pathway;
KW invasive growth pathway; biological pathway; ss.
XX OS
OS Saccharomyces cerevisiae.
XX FH
FH Key Location/Qualifiers
FT CDS 1001..1525
FT /*tag= a
FT /product= "ergosterol-pathway polypeptide"
XX WO200058520-A1.
FN
XX
XX
PD 05-OCT-2000.
XX
XX 29-MAR-2000; 2000WO-US008555.
XX PF
XX 31-MAR-1999; 99US-00282243.
XX PR
XX (ROSE-) ROSETTA INPHARMATICS INC.
XX PA
XX Roberts CJ;
XX PI
XX
DR WPI; 2000-664929/64.
DR P-PSDB; AAB19091.
XX
XX Identifying reporter and target genes for particular biological pathway
PT of interest, useful for drug designing, involves identifying a gene which
PT clusters to geneset associated with biological pathway.
XX
PS Claim 45; Fig 8; 239pp; English.
XX
CC The present sequence represents a reporter gene from the ergosterol
CC pathway. The specification also describes reporter genes from the protein
CC kinase C pathway, and the invasive growth pathway. The genes are
CC identified using methods of the invention. The specification describes
CC methods for identifying a reporter gene or a target gene for a particular
CC biological pathway in a cell. The methods comprise identifying a gene
CC which clusters to a geneset associated with the pathway. The gene
CC identified is necessary for normal function of the pathway. The methods
CC are useful for analysing the activity of biological pathways and in drug
CC designing, drug therapies, or other biological agents e.g. insecticides,
CC herbicides, fungicides, antibodies or antivirals. The reporter gene is
CC useful as a specific target for drugs which are designed to enhance,
CC inhibit or modulate a particular biological pathway
XX
SQ Sequence 1725 BP; 469 A; 315 C; 365 G; 576 T; 0 U; 0 Other;

Query Match 91.6%; Score 17.4; DB 3; Length 1725;
Best Local Similarity 94.7%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
Db 386 CTTTACTTCATAGCTTTG 404

RESULT 42
ADF70573/C
ID ADF70573 standard; DNA; 1788 BP.
XX AC ADF70573;
XX AC
DT 12-FEB-2004 (first entry)
XX
XX Orphan receptor ligand-related human protein gene SeqID196.
XX ligand; orphan receptor protein; fusion protein; fluorescent protein;
KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.
XX
XX Homo sapiens.
XX OS
XX WO2003071272-A1.
FN
XX 28-AUG-2003.
PD
XX 21-FEB-2003; 2003WO-JP001901.
PF
XX 22-FEB-2002; 2002JP-00045728.
PR
XX 23-JUL-2002; 2002JP-00213949.
PR
XX 11-OCT-2002; 2002JP-00298237.
PR
XX (TAKE) TAKEDA CHEM IND LTD.
FA
XX Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
PI WPI; 2003-697654/66.
XX P-PSDB; ADF70471.
DR
XX Transformation of cells with a fusion protein of an orphan receptor
PT protein with a fluorescent protein useful for identification of ligands
PT to the orphan receptor.
XX
XX Example 4; SEQ ID NO 196; 594pp; Japanese.
PS
XX This invention relates to a novel method of identifying ligands to an
CC orphan receptor protein which comprises transforming cells with DNA
CC encoding a fusion protein of the orphan receptor with a fluorescent
CC protein, so that the fusion protein is expressed in the cells (or cell
CC membranes isolated from them) and contacting the cells with the potential
CC ligand to be tested. A suitable fluorescent protein for incorporation in
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
CC identification of ligands binding to an orphan receptor protein.
XX
SQ Sequence 1788 BP; 444 A; 473 C; 412 G; 459 T; 0 U; 0 Other;

Query Match 91.6%; Score 17.4; DB 10; Length 1788;
Best Local Similarity 94.7%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
Db 1059 CTTTACTTCATAGCTTTG 1041

RESULT 43
AEF80156
ID AEF80156 standard; DNA; 188273 BP.
XX AC AEF80156;
XX AC
DT 06-APR-2006 (first entry)
XX
XX Cancer-associated polypeptide genomic DNA sequence hb28-002 SEQ ID NO:52.
DE ds; gene; DNA microarray; cancer; neoplasm; cytostatic; diagnosis.
KW
XX

```
OS Homo sapiens.
XX US2006024677-A1.
PN
XX
XX
PD 02-FEB-2006.
XX
XX 20-JUL-2004; 2004US-00895974.
XX
XX 20-JUL-2004; 2004US-00895974.
XX
XX (MORE/) MORIS D W
PA (MALA/) MALANDRO M S.
PA (LAIA/) LAI A.
PA (TSEC/) TSE C.
PA (PAT/) FATTAY A.
XX
XX Morris DW, Malandro MS, Lai A, Tse C, Fattay A;
XX
XX WPI; 2006-135411/14.
DR P-PSDB; AEF80158.
XX
XX Nucleic acid array for detecting cancer-associated (CA) nucleic acid,
PT consists of nucleic acid probes having specific contiguous nucleotides of
PT CA polynucleotide.
XX
XX Disclosure; SEQ ID NO 52; 264pp; English.
XX
XX The invention relates to a novel nucleic acid array (I) for detecting a
CC cancer-associated (CA) nucleic acid, consisting of 2 or more nucleic acid
CC probes each comprising 10 or more contiguous nucleotides of 2 or more CA
CC polynucleotide sequences, or its complement. The invention has cytostatic
CC activity. The nucleic acid array is useful for detecting a CA nucleic
CC acid. An antibody of the invention is useful for detecting the presence
CC or absence of cancer cells. A method of the invention is useful for
CC inhibiting expression of a CA gene in a cell, or for treating cancer. The
CC CA polynucleotide or polypeptide as mentioned in (I) or (II) is useful as
CC vaccine for treating cancer e.g. lymphoma or leukemia. The present
CC sequence represents a human genomic sequence encoding a CA polypeptide
CC (CAP) of the invention.
XX
XX Sequence 188273 BP; 52885 A; 38491 C; 40162 G; 55953 T; 0 U; 782 Other;
SQ
Query Match 91.6%; Score 17.4; DB 15; Length 188273;
Best Local Similarity 94.7%; Pred. No. 4.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGCTTTG 19
DB 53975 CATTACTTCATAGCTTTG 53993
RESULT 44
AEE07706
ID AEE07706 standard; cDNA; 168 BP.
AC AEE07706;
XX
XX 26-JAN-2006 (first entry)
XX
XX Tobacco p450 cDNA SEQ ID NO: 1152.
XX
XX Plant breeding; plant; ss; coding sequence; cytochrome p450;
KW transgenic plant; senescence.
XX
XX Nicotiana tabacum.
OS
XX WO200511217-A2.
PN
XX
XX 24-NOV-2005.
PD
XX
XX 27-APR-2005; 2005WO-US014803.
XX
XX 29-APR-2004; 2004US-0566235P.
XX
03-SEP-2004; 2004US-00934944.
PR
PR 03-SEP-2004; 2004US-0607357P.
PR
PR 17-SEP-2004; 2004US-00943507.
PR
PR 15-OCT-2004; 2004WO-US034065.
PR
PR 15-OCT-2004; 2004WO-US034218.
PR
PR 25-JAN-2005; 2005US-0646764P.
PR
PR 24-MAR-2005; 2005US-0665097P.
PR
PR 24-MAR-2005; 2005US-0665451P.
XX
XX (USSM-) US SMOKELESS TOBACCO CO.
PA
XX
XX Xu D, Nielsen MT;
PI
XX
XX WPI; 2005-786788/80.
DR
XX
XX Producing a tobacco plant having decreased expression of a nicotine
PT demethylase gene comprises crossing a first tobacco plant with a second
PT tobacco plant and germinating the collected seed of an F1 progeny plant.
XX
XX Claim 32; SEQ ID NO 1152; 641pp; English.
XX
XX The invention relates to a breeding method for producing a tobacco plant
CC with reduced expression of a nicotine demethylase gene comprises crossing
CC a first tobacco plant with variant nicotine demethylase gene expression
CC with a second tobacco plant with at least one phenotypic trait to produce
CC an F1 progeny plant, the seed of which is collected and germinated to
CC produce a tobacco plant having reduced expression of a nicotine
CC demethylase gene. Also included are breeding a nicotine demethylase
CC deficiency trait into a tobacco plant, producing a tobacco seed,
CC developing a tobacco plant in a tobacco breeding program, a tissue
CC culture of regenerable tobacco cells obtained from the tobacco plant of
CC the invention, producing a tobacco product, a breeding method for
CC producing a tobacco plant having a modified attribute, a method of
CC breeding an attribute into a tobacco plant, a tobacco plant or its
CC components produced by the method of breeding a nicotine demethylase
CC deficiency trait into a tobacco plant, producing a tobacco seed, producing
CC a tobacco plant having a modified attribute or developing a tobacco plant
CC in a tobacco breeding program, an isolated genetic marker comprising a
CC nucleic acid sequence that is substantially identical to a nucleic acid
CC sequence given in the specification (the nucleic acids comprise isolated
CC cytochrome p450 cDNAs), an expression vector comprising the isolated nucleic acid
CC sequence, a plant or plant component comprising the isolated nucleic acid
CC sequence, a plant produced from a germinated seed of the plant, reducing
CC the expression or enzymatic activity of a constitutive, or an ethylene
CC induced or senescence induced tobacco polypeptide in a plant cell, and
CC increasing the expression or enzymatic activity of a constitutive, or an
CC ethylene or senescence induced tobacco polypeptide in a plant cell. The
CC phenotypic trait comprises disease resistance, high yield, high grade
CC index, curability, curing quality, mechanical harvestability, holding
CC ability, leaf quality, height, maturation, stalk size, or leaf number per
CC plant. The breeding method for producing a tobacco plant having decreased
CC expression of a nicotine demethylase gene is useful developing desirable
CC (non-genetically engineered) germplasm. The plant is useful in producing
CC (smokeless) tobacco products. The tobacco product is a moist or dry
CC snuff, a chewing tobacco, a cigarette product, a cigar product, a
CC cigarillo, a pipe tobacco, or bidis. The p450 cDNAs were isolated using
CC degenerate PCR primers designed against cytochrome p450 motifs. The
CC present sequence is a cytochrome p450 cDNA of the invention.
XX
XX Sequence 168 BP; 44 A; 23 C; 34 G; 67 T; 0 U; 0 Other;
SQ
Query Match 86.3%; Score 16.4; DB 14; Length 168;
Best Local Similarity 94.4%; Pred. No. 7.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TTTTACTTCATAGCTTTG 19
DB 35 TTTTACTTCATAGCTTTG 52
RESULT 45
AEF98099
ID AEF98099 standard; cDNA; 168 BP.
```


CC acid molecules, the method provides information on the more interesting
CC and relevant internal portions, such as ORFs. The method also permits the
CC construction of contigs of sequenced nucleic acid molecules

XX SQ Sequence 354 BP; 118 A; 73 C; 80 G; 79 T; 0 U; 4 Other;
Query Match 86.3%; Score 16.4; DB 3; Length 354;
Best Local Similarity 94.4%; Pred. No. 7.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCAGAGTCTTTG 19
|||||

Db 175 TTTACTTCAGAGTCTTTG 158

RESULT 47

ABT12676/c

ID ABT12676 standard; DNA; 354 BP.

XX AC ABT12676;

XX DT 02-JAN-2003 (first entry)

XX DE Orestes sequencing method related DNA SEQ ID No 382.

XX KW Open reading frame; ORF; low stringency; sequencing; ORESTES; bronchitis;

XX KW identification; internal region; population study; heart disease;

XX KW Alzheimer's disease; autoimmune disease; congenital disease; human; ds.

XX OS Homo sapiens.

XX XX WO200274994-A2.

XX XX 26-SEP-2002.

XX XX 01-NOV-2001; 2001WO-US046665.

XX XX 07-NOV-2000; 2000US-0246313P.

XX XX (LUDW-) LUDWIG INST CANCER RES.

XX PI Simpson AJG, Dias-Neto E, Brentani RR;

XX WPI; 2002-750567/81.

XX XX Determining ORFs of the genome by contacting mRNA with a single

XX PT oligonucleotide primer, useful for population studies that identify genes

XX PT associated with diseases such as cardiovascular, autoimmune and

XX PT congenital diseases.

XX XX Example 7; Page 160; 166pp; English.

XX XX The invention relates to a novel method for determining open reading

XX CC frames (ORFs) of the genome of an organism comprising contacting

XX CC messenger RNA from a cell of the organism with a single, oligonucleotide

XX CC primer at low stringency. The method is an improved ORESTES sequencing

XX CC method. The methods of the present invention are useful for sequencing

XX CC of, essentially, the entire coding component of an organism,

XX CC identification of previously unknown nucleic acid molecules, assigning

XX CC previously identified nucleotide sequences to internal regions of genes

XX CC and population studies that identify genes associated with diseases such

XX CC as heart disease, bronchitis, Alzheimer's disease, autoimmune diseases

XX CC and congenital diseases. This polynucleotide represents a DNA sequence

XX CC that is used in the novel method of the invention

XX SQ Sequence 354 BP; 118 A; 73 C; 80 G; 79 T; 0 U; 4 Other;

Query Match 86.3%; Score 16.4; DB 6; Length 354;

Best Local Similarity 94.4%; Pred. No. 7.8e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCAGAGTCTTTG 19

|||||

Db 175 TTTACTTCAGAGTCTTTG 158

RESULT 48

ACD91970/c

XX ID ACD91970 standard; cDNA; 354 BP.

XX AC ACD91970;

XX DT 23-SEP-2003 (first entry)

XX DE Human colon cancer cell expressed cDNA #382.

XX KW Open reading frame detection; genome sequencing; colon cancer;

XX KW breast cancer; population genome analysis; genetic shift; cancer;

XX KW antibiotic resistance; antibiotic non-tolerance; congenital disease;

XX KW agriculture; food crop genome; resistance gene; retrovirus;

XX KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;

XX KW gene; ss.

XX OS Homo sapiens.

XX PN US2002155438-A1.

XX XX 24-OCT-2002.

XX XX 27-SEP-1999; 99US-00406117.

XX XX 20-NOV-1998; 98US-00196716.

XX XX (SIMP/) SIMPSON A J G.

XX XX (NETO/) NETO E D.

XX XX (BREN/) BRENTANI R R.

XX PI Simpson AJG, Neto ED, Brentani RR;

XX WPI; 2003-182626/18.

XX XX Determining open reading frames of genome of an organism e.g. a human

XX PT suffering from cancer involves use of single oligonucleotide primer at

XX PT low stringency for preparing single-stranded cDNA from mRNA of

XX PT individual.

XX XX Example 7; Page 56; 959pp; English.

XX XX The invention describes a method of determining open reading frames in

XX CC the genome of organism, comprising contacting mRNA from cell of organism

XX CC with a single oligonucleotide primer (I) at low stringency, preparing

XX CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying

XX CC cDNA, sequencing the product, and repeating the contacting, preparing

XX CC and amplifying steps with different primers and sequencing resulting

XX CC nucleic acids. The method is useful for: determining that a known

XX CC nucleotide sequence of a genome of an organism corresponds to a

XX CC nucleic acid molecule from a genome of an organism; and for sequencing

XX CC all or part of a genome of an organism. mRNA is obtained from mammalian

XX CC or human cell which is associated with a pathological condition e.g. a

XX CC colon cancer or breast cancer cell. The method is useful for analyses of

XX CC populations of subjects and can be used to carry out genetic analyses of

XX CC large or small populations. further, it can be used to study living

XX CC systems to determine if, e.g. there have been genetic shifts which render

XX CC an individual or population more or less likely to be afflicted with

XX CC diseases such as cancer, to determine antibiotic resistance or non-

XX CC tolerance, and so forth. The method can also be used in the study of

XX CC congenital diseases, and the risk of affliction to a fetus, as well as

XX CC the study of whether the conditions are likely to be passed to offspring

XX CC through ova or sperm. The analyses for pathological conditions can be

XX CC carried out in all animals, plants, birds, fish, etc. Using this method,

XX CC in the area of agriculture, for example the genomes of food crops can be

XX CC studied to determine if resistance genes are present, defects in plant

XX CC genomes can also be studied in this way. Similarly, the method permits

XX CC determination of the pathogens which integrate into the genome, such as

XX CC retroviruses and other integrating viruses such as influenza virus, have

CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX
SQ Sequence 354 BP; 118 A; 73 C; 80 G; 79 T; 0 U; 4 Other;

Query Match 86.3%; Score 16.4; DB 10; Length 354;
Best Local Similarity 94.4%; Pred. No. 7.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19
|||||
Db 175 TTTACTTCAGAGTCTTTG 158

RESULT 49
AAC28686/c
ID AAC28686 standard; cDNA; 407 BP.

XX AAC28686;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 32761.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 32761; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors

XX Sequence 407 BP; 124 A; 78 C; 81 G; 123 T; 0 U; 1 Other;

Query Match 86.3%; Score 16.4; DB 3; Length 407;
Best Local Similarity 94.4%; Pred. No. 7.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19
|||||
Db 40 TTTACTTCATAGTCTTTG 23

RESULT 50

AAK74755/c

ID AAK74755 standard; DNA; 429 BP.

XX AAK74755;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29567.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 30-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 11-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

PN WO200155314-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001324.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254057P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX
PI

```
XX WPI; 2001-502630/55.
XX P-PSDB; AAM92695.
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX Claim 1; SEQ ID NO 784; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a cDNA encoding a digestive
XX system antigen of the invention
XX
XX Sequence 491 BP; 144 A; 90 C; 68 G; 186 T; 0 U; 3 Other;
SQ
Query Match 84.2%; Score 16; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TTACTTCATAGTCTTT 18
Db 312 TTACTTCATAGTCTTT 327
RESULT 52
AAI97742
ID AAI97742 standard; cDNA; 910 BP.
XX
XX AAI97742;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human neuroblastoma expressed polynucleotide SEQ ID NO 3817.
XX
XX Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
XX Homo sapiens.
XX
XX WO200166719-A1.
XX
XX 13-SEP-2001.
XX
XX 02-MAR-2001; 2001WO-JP001629.
XX
XX 07-MAR-2000; 2000JP-00159195.
XX
XX (CHIB-) CHIBA PREFECTURE.
XX
XX (HISM ) HISAMITSU PHARM CO LTD.
XX
XX Nakagawara A;
XX
XX WPI; 2001-565584/63.
XX
XX Nucleic acids originating in gene expressed in human neuroblastoma,
XX useful as probe or primer in diagnosing prognosis of human neuroblastoma,
XX malignancy and susceptibility indicator or tumor marker for anti-cancer
XX agents.
XX
XX Claim 1; Page 2803; 2979pp; Japanese.
XX
XX The invention relates to novel genes (AAI93926-AAI97963) expressed in
XX human neuroblastoma. The nucleic acids are applicable as a probe or
XX primer in diagnosing the prognosis of human neuroblastoma, malignancy and
XX susceptibility indicators or tumour markers for anti-cancer agents. The
XX gene information for diagnosing prognosis is related to factors similar
XX to that for N-myc and TrkA genes
XX
XX Sequence 910 BP; 203 A; 219 C; 226 G; 232 T; 0 U; 30 Other;
SQ
Query Match 84.2%; Score 16; DB 4; Length 910;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 TTTACTTCATAGTCTT 17
Db 184 TTTACTTCATAGTCTT 199
RESULT 53
AAI97741/C
ID AAI97741 standard; cDNA; 981 BP.
XX
XX AAI97741;
XX
XX 13-NOV-2001 (first entry)
XX
XX Human neuroblastoma expressed polynucleotide SEQ ID NO 3816.
XX
XX Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
XX Homo sapiens.
XX
XX WO200166719-A1.
XX
XX 13-SEP-2001.
XX
XX 02-MAR-2001; 2001WO-JP001629.
XX
XX 07-MAR-2000; 2000JP-00159195.
XX
XX (CHIB-) CHIBA PREFECTURE.
XX
XX (HISM ) HISAMITSU PHARM CO LTD.
XX
XX Nakagawara A;
XX
XX WPI; 2001-565584/63.
XX
XX Nucleic acids originating in gene expressed in human neuroblastoma,
XX useful as probe or primer in diagnosing prognosis of human neuroblastoma,
XX malignancy and susceptibility indicator or tumor marker for anti-cancer
XX agents.
XX
XX Claim 1; Page 2802; 2979pp; Japanese.
XX
XX The invention relates to novel genes (AAI93926-AAI97963) expressed in
XX human neuroblastoma. The nucleic acids are applicable as a probe or
XX primer in diagnosing the prognosis of human neuroblastoma, malignancy and
XX susceptibility indicators or tumour markers for anti-cancer agents. The
XX gene information for diagnosing prognosis is related to factors similar
XX to that for N-myc and TrkA genes
XX
XX Sequence 981 BP; 261 A; 226 C; 203 G; 240 T; 0 U; 51 Other;
SQ
Query Match 84.2%; Score 16; DB 4; Length 981;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 TTTACTTCATAGTCTT 17
Db 298 TTTACTTCATAGTCTT 283
RESULT 54
AAK90456
ID AAK90456 standard; DNA; 1596 BP.
XX
XX AAK90456;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human digestive system antigen genomic sequence SEQ ID NO: 4032.
XX
```

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX Homo sapiens.
XX WO200155314-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US0011324.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

```
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 4032; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
XX Sequence 1596 BP; 486 A; 304 C; 259 G; 547 T; 0 U; 0 Other;
SQ
Query Match 84.2%; Score 16; DB 4; Length 1596;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 3 TTACTTCATAGTCTTT 18
Db 1289 TTACTTCATAGTCTTT 1304
RESULT 55
AAK90457
ID AAK90457 standard; DNA; 1596 BP.
XX
XX AAK90457;
XX
XX 05-NOV-2001 (first entry)
DT
XX
XX Human digestive system antigen genomic sequence SEQ ID NO: 4033.
DE
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
XX Homo sapiens.
XX
XX WO200155314-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001324.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 03-SEP-2000; 2000US-0229509P.
PR 03-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
```


PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246479P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Baraash SC, Ruben SM;

XX WPI; 2001-502630/55.

XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.

PS Disclosure; SEQ ID NO 4033; 986pp; English.

XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention

SQ Sequence 1596 BP; 486 A; 304 C; 259 G; 547 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 4; Length 1596;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTACTTCATAGCTTTT 18
Db 1289 TTACTTCATAGCTTTT 1304

RESULT 56

AEF99265
ID AEF99265 standard; DNA; 3447 BP.

AC AEF99265;

XX 20-APR-2006 (first entry)

XX Human EFHC1 splice variant DNA.

XX neuron; EFHC1; epilepsy; therapeutic; pharmaceutical; cell death;
KW apoptosis; diagnostic; neurological disease; Anticonvulsant; ds.

XX Homo sapiens.

XX WO2006019978-A2.

XX 23-FEB-2006.

XX 15-JUL-2005; 2005WO-US025093.

XX 16-JUL-2004; 2004US-0588769P.

XX (REGC) UNIV CALIFORNIA.

PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.

PA (NAME-) INST NACIONAL NEUROLOGIA & NEUROCIROGIA.

PA (UYHO-) UNIV HONDURAS NACIONAL AUTONOMA.

XX Delgado-Escueta AV, Yamakawa K, Suzuki T, Medina-Hernandez MT;
PI Villatela MEA;

XX WPI; 2006-184370/19.

XX EMBL; AL22084.

XX Inhibiting growth of a neuronal cell for treating epilepsy, comprises
PT contacting a cell with EFHC1 or its analog, or an EFHC1 agonist.

XX Example 11; SEQ ID NO 8; 131pp; English.

XX This invention describes a novel of inhibiting growth of a neuronal cell
CC and comprises contacting the cell with EFHC1 or its analog, or an EFHC1
CC agonist e.g. an R-type voltage-dependent Ca²⁺ channel. The invention also
CC describes: a) a method of potentiating an epilepsy therapy in a patient
CC by administering EFHC1 or its analog, or EFHC1 agonist, and administering
CC an anti-epileptic drug; b) a method of identifying agents that promote
CC cell death in a mammalian cell; c) a method of identifying an agent that
CC modulates activity of an EFHC1 polypeptide; d) a method of identifying a
CC agent that modulates activity of an apoptotic drug by providing a
CC mammalian cell or mammalian cell culture engineered to overexpress EFHC1;
CC e) a method of monitoring the efficacy of an epilepsy treatment by
CC detecting levels of AEF99260 or AEF99261 where an increased level of the
CC EFHC1 polypeptide or EFHC1 polynucleotide or gene copy number in the
CC biological sample compared to the level or copy number in a biological
CC sample from the patient before the treatment is indicative of efficacious
CC treatment; f) a method of diagnosing epilepsy in a mammalian subject; g)
CC a vector or host cell containing a nucleic acid encoding an EFHC1
CC polypeptide; h) an antibody that specifically binds to an EFHC1
CC polypeptide; and i) a diagnostic kit comprising a solid support
CC (nitrocellulose, latex or plastic material), a detection reagent (anti-
CC immunoglobulin, protein G, protein A or lectin) and a reporter group is
CC chosen from radioisotopes, fluorescent groups, luminescent groups,
CC enzymes, biotin and dye particles. the anti-epileptic drug promotes
CC apoptosis and is chosen from carbamazepine, ethosuximide, CBZ epoxide,
CC lamotrigine, oxcarbazepine, sodium valproate, acetazolamide, clobazam,
CC clonazepam, gabapentin, levetiracetam, phenobarbital, phenytoin,
CC piracetam, primidone, tiagabine, topiramate, zonisamide or vigabatrin.
CC The methods of the invention are used to a) inhibit growth of a neuronal

CC cell; b) potentiate an epilepsy therapy in a patient; c) potentiate
CC treatment for juvenile myoclonic epilepsy (JMC); d) monitor the efficacy
CC of epilepsy treatment and e) enables improvement of epilepsy treatment
CC without causing substantial toxicity. This sequence represents a splice
CC variant of the human BFHCl gene used in the invention.

XX SQ Sequence 3447 BP; 1022 A; 709 C; 677 G; 1039 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 15; Length 3447;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTT 18

|||||

Db 1967 TTACTTCATAGTCTTT 1982

RESULT 57

ADJ80212/c

ID ADJ80212 standard; cDNA; 3737 BP.

XX

AC ADJ80212;

XX

DT 06-MAY-2004 (first entry)

XX

DE Novel human nucleic acid-associated protein coding sequence #30.

XX

KW ds; gene; cytostatic; antiarteriosclerotic; cerebroprotective;

KW antiparkinsonian; anticonvulsant; anti-HIV; anti-allergic; antibacterial;

KW virucide; gene therapy; nucleic acid-associated protein; cancer;

KW atherosclerosis; stroke; Parkinson's disease; epilepsy;

KW Cushing's syndrome; AIDS; allergy; microarray element;

KW protein-protein interaction; drug-target interaction; gene expression;

KW chromosomal mapping; diagnosis.

XX

OS Homo sapiens.

XX

PN W02003038052-A2.

XX

PD 08-MAY-2003.

XX

PF 29-OCT-2002; 2002WO-US034846.

XX

PR 29-OCT-2001; 2001US-0348442P.

PR

PR 01-NOV-2001; 2001US-0335544P.

PR

PR 05-NOV-2001; 2001US-0337535P.

PR

PR 09-NOV-2001; 2001US-0344650P.

PR

PR 15-NOV-2001; 2001US-0334762P.

XX

XX (INCY-) INCYTE GENOMICS INC.

PA

XX Becha SD, Borowsky ML, Burford N, Chawla NK, Elliott VS;

XX Emerling BM, Forsythe IJ, Gietzen KJ, Gorvad AE, Griffin JA;

XX Hafalia AJA, Ison CH, Lal PG, Lee EA, Lee S, Lee SY, Marquis JP;

XX Ramkumar J, Sprague WM, Swarnakar A, Tang YT, Warren BA, Yang J;

XX Yue H, Zebardjian Y;

XX WPI; 2003-430514/40.

DR

DR P-PSDB; ADJ80154.

XX

PT New human nucleic acid-associated protein (NAAP) and polynucleotide,

PT useful for diagnosing, treating, and preventing disorders associated with

PT aberrant expression of NAAP, e.g. cancer, AIDS, stroke or infection.

XX

XX Claim 5; SEQ ID NO 88; 443pp; English.

XX

CC The invention relates to novel human nucleic acid-associated proteins and

CC genes encoding them, sequences that have at least 90-99 % identity to the

CC sequences; or biologically active or immunogenic fragments of these. The

CC polypeptides and polynucleotides are useful in diagnosing, treating and

CC preventing disorders associated with aberrant expression of NAAP, such as

CC cell proliferative (e.g. cancer or atherosclerosis), neurological (e.g.

CC syndrome), autoimmune/inflammatory (e.g. AIDS or allergies), or
CC infections. These may also be used as elements on a microarray which may
CC monitor or measure protein-protein interactions, drug-target
CC interactions, and gene expression profiles. The polynucleotide may also
CC be used in chromosomal mapping and in various diagnostic assays. These
CC are also useful in assessing the effects of exogenous compounds on the
CC expression of nucleic acids and amino acid sequences of NAAP, in
CC facilitating drug discovery process, and in investigating the
CC pathogenesis of diseases or medical conditions. This sequence corresponds
CC to the gene encoding one of the proteins of the inventions.

XX SQ Sequence 3737 BP; 1046 A; 1017 C; 1018 G; 656 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 10; Length 3737;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTACTTCATAGTCTT 17

|||||

Db 844 TTACTTCATAGTCTT 829

RESULT 58

AEC20314/c

ID AEC20314 standard; DNA; 3737 BP.

XX

AC AEC20314;

XX

DT 03-NOV-2005 (first entry)

XX

DE Human nucleic acid-associated protein (NAAP) coding sequence - SEQ ID 88.

XX

KW nucleic acid-associated protein; hyperproliferation; atherosclerosis;

KW antiarteriosclerotic; hepatitis; antiinflammatory; hepatotropic;

KW psoriasis; antipsoriatic; cancer; cytostatic; neurological disease;

KW neuroprotective; epilepsy; anticonvulsant; cerebrovascular ischemia;

KW cerebroprotective; vasotropic; dementia; neurotropic; parkinsons disease;

KW antiparkinsonian; multiple sclerosis; anemia; antianemic;

KW Cushing's disease; endocrine-gen.; spina bifida; autoimmune disease;

KW immunosuppressive; inflammation; acquired immune deficiency syndrome;

KW anti-hiv; allergy; anti-allergic; contact dermatitis; dermatological;

KW diabetes; antidiabetic; glomerulonephritis; nephrotropic; osteoporosis;

KW osteopathic; osteoarthritis; antiarthritic; rheumatoid arthritis;

KW antineumatic; bacterial infection; antibacterial; viral infection;

KW virucide; fungal infection; fungicide; parasitic infection;

KW antiparasitic; gene; ds.

XX

OS Homo sapiens.

XX

PN US2005186569-A1.

XX

PD 25-AUG-2005.

XX

PF 31-MAR-2004; 2004US-00491472.

XX

PR 29-OCT-2001; 2001US-0348443P.

PR

PR 01-NOV-2001; 2001US-0335544P.

PR

PR 05-NOV-2001; 2001US-0337535P.

PR

PR 09-NOV-2001; 2001US-0344650P.

PR

PR 15-NOV-2001; 2001US-0334762P.

PR

PR 29-OCT-2002; 2002WO-US034846.

XX

XX (BECH/) BECHA S D.

XX (BORO/) BOROWSKY M L.

XX (BURF/) BURFORD N.

XX (CHAW/) CHAWLA N K.

XX (ELLI/) ELLIOTT V S.

XX (EMER/) EMERLING B M.

XX (FORS/) FORSYTHE I J.

XX (GIET/) GIETZEN K J.

XX (GORV/) GORVAD A E.

XX (GRIF/) GRIFFIN J A.

XX (HAFa/) HAFALIA A J A.

PA (ISON/) ISON C H.
 PA (LALP/) LAL P.
 PA (LEEE/) LEE E A.
 PA (LEES/) LEE S.
 PA (LEES/) LEE S Y.
 PA (MARQ/) MARQUIS J P.
 PA (RAMK/) RAMKUMAR J.
 PA (SPRA/) SPRAGUE W W W.
 PA (SWAR/) SWARNAKAR A.
 PA (TANG/) TANG Y T.
 PA (WARR/) WARREN B A.
 PA (YANG/) YANG J.
 PA (YUEH/) YUE H.
 PA (ZEBAR/) ZEBARJADIAN Y.
 XX
 PI Becha SD, Borowsky ML, Burford N, Chawla NK, Elliott VS;
 PI Emiling BM, Forsythe IJ, Gletzen KJ, Gorvad AE, Griffin JA;
 PI Hafalia AJA, Lee CH, Lal P, Lee EA, Lee S, Lee SY, Marquis JP;
 PI Ramkumar J, Sprague WW, Swarnakar A, Tang YT, Warren BA, Yang J;
 PI Yue H, Zebbarjadian Y;
 XX
 DR WPI: 2005-570800/58.
 DR P-PSDB; AEC20256.
 XX
 PT New isolated nucleic acid-associated proteins (NAAP) and polynucleotides,
 PT useful for diagnosing, treating, or preventing cell proliferative, or
 PT neurological, developmental, or autoimmune/inflammatory disorders, or
 PT infections.
 XX
 PS Claim 12; SEQ ID NO 88; 278pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
 CC the invention are useful for diagnosing, treating, or preventing
 CC disorders associated with aberrant expression of NAAP. Specifically, the
 CC disorders include cell proliferative disorders (e.g. atherosclerosis,
 CC hepatitis, psoriasis, and cancer), neurological disorders (e.g. epilepsy,
 CC stroke, dementia, Parkinson's disease, and multiple sclerosis),
 CC developmental disorders (e.g. anemia, Cushing's syndrome, and spina
 CC bifida), autoimmune/inflammatory disorders (e.g. AIDS, allergy, contact
 CC dermatitis, diabetes, glomerulonephritis, osteoporosis, osteoarthritis,
 CC and rheumatoid arthritis), and infections (e.g. bacterial, viral, fungal,
 CC and parasitic infections). The present DNA sequence encodes a human NAAP
 CC of the invention.
 XX
 SQ Sequence 3737 BP; 1046 A; 1017 C; 1018 G; 656 T; 0 U; 0 Other;
 Query Match 84.2%; Score 16; DB 14; Length 3737;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 TTACTTCATAGTCTT 17
 Db 844 TTACTTCATAGTCTT 829
 RESULT 59
 AAC75928
 ID AAC75928 standard; cDNA; 3739 BP.
 XX
 AC AAC75928;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1483 polynucleotide sequence SEQ ID NO:2965.
 XX
 KW Human; open reading frame; ORFX: detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200058473-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX
 XX 31-MAR-2000; 2000WO-US008621.
 PP
 XX 31-MAR-1999; 99US-0127607P.
 PR
 XX 02-APR-1999; 99US-0127636P.
 PR
 XX 05-APR-1999; 99US-0127728P.
 PR
 XX 30-MAR-2000; 2000US-00540763.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkets RA, Leach M;
 PI
 XX WPI: 2000-602362/57.
 DR
 XX P-PSDB; AAB41719.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 5; Page 2200-2202; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 3739 BP; 1078 A; 787 C; 764 G; 1110 T; 0 U; 0 Other;
 Query Match 84.2%; Score 16; DB 3; Length 3739;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 TTACTTCATAGTCTT 18
 Db 2259 TTACTTCATAGTCTT 2274
 RESULT 60
 AEF22147/c
 ID AEF22147 standard; DNA; 3786 BP.
 XX
 AC AEF22147;
 XX
 DT 09-MAR-2006 (first entry)

XX DE Human basal transcription modulator splice variant DNA #12.
XX KW Cytostatic; diagnosis; transcription factor; neoplasm; ds.
XX OS Homo sapiens.
XX PN WO2006005042-A2.
XX PD 12-JAN-2006.
XX PF 30-JUN-2005; 2005WO-US023708.
XX PR 30-JUN-2004; 2004US-0584784P.
XX PA (CEMI-) CEMINES INC.
XX PI Shen D, Neuman T, Palm K;
XX WPI; 2006-100776/10.

XX DR Diagnosing cancer, useful for treating cancer, e.g. lung, breast,
XX PT prostate, skin, gastrointestinal cancer, comprises determining the
XX PT expression of at least one splice variant of each of the basal
XX PT transcription factors.
XX PS Disclosure; Fig 9; 131pp; English.
XX CC The invention relates to a method of diagnosing cancer which comprises
XX CC determining the expression of at least one splice variant of each of the
XX CC basal transcription factors, where expression of each of the basal
XX CC transcription factor splice variants is distinguished from expression of
XX CC its wildtype isoform, and where the expression pattern of the basal
XX CC transcription factor splice variants is indicative of cancer. The methods
XX CC and bioactive agents are useful for treating cancer, e.g. lung cancer,
XX CC gastrointestinal cancer, breast cancer, prostate cancer, skin cancer,
XX CC sarcoma, endocrine cancer, neural cancer, bladder cancer, cervical
XX CC cancer, renal cancer and hematopoietic cancer. The present sequence
XX CC represents a human basal transcription modulator splice variant DNA.

SQ Sequence 3786 BP; 1063 A; 1031 C; 1030 G; 662 T; 0 U; 0 Other;
Query Match 84.2%; Score 16; DB 15; Length 3786;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 TTACTTCATAGTCTT 17
DB 811 TTACTTCATAGTCTT 796

RESULT 61
ACN39480/c
ID ACN39480 standard; CDNA; 4816 BP.
XX AC ACN39480;
XX 18-NOV-2004 (first entry)
XX DE Tumour-associated antigenic target (TAT) cDNA DNA325691, SEQ ID NO:3645.
XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX KW tumour; diagnosis; cell proliferative disorder; breast cancer;
XX KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX KW central nervous system cancer; bladder cancer; pancreatic cancer;
XX KW cervical cancer; melanoma; leukaemia; hybridisation probe;
XX KW chromosome identification; chromosome mapping; gene mapping;
XX KW gene therapy; cytostatic; gene; ss.
XX OS Homo sapiens.
XX PN WO2004030615-A2.
XX PR 16-DEC-2003; 2003US-00737082.

PD 15-APR-2004.
XX 29-SEP-2003; 2003WO-US028547.
XX 02-OCT-2002; 2002US-0414971P.
XX PA (GETH) GEMENTECH INC.
XX PI Wu TD, Zhang Z, Zhou Y;
XX WPI; 2004-347921/32.
XX DR New tumor-associated antigenic target polypeptides and nucleic acids,
XX PT useful in preparing a medicament for treating or detecting a
XX PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX PT prostate cancer or tumor.

XX PS Claim 1; SEQ ID NO 3645; 7273pp; English.

XX CC The invention relates to human tumour-associated antigenic target (TAT)
XX CC polypeptides, and their related nucleic acids. The TAT polypeptides are
XX CC overexpressed in cancer tissues compared to normal tissues, and may thus
XX CC serve as effective targets for the diagnosis and treatment of cancer in
XX CC mammals. The invention also relates to nucleic acid and polypeptide
XX CC sequences at least 80% identical to the TAT nucleic acids and
XX CC polypeptides, expression vectors and host cells comprising a TAT nucleic
XX CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX CC TAT polypeptide; and methods and compositions for the treatment or
XX CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX CC antibodies, antagonists, binding molecules and compositions are useful
XX CC for diagnosing or treating a cell proliferative disorder associated with
XX CC increased TAT expression, particularly cancers such as breast cancer,
XX CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX CC cancer, pancreatic cancer, cervical cancer, cancers of the central
XX CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX CC used as hybridisation probes, in chromosome and gene mapping, in
XX CC chromosome identification and in gene therapy. The present sequence
XX CC represents a TAT nucleic acid of the invention

SQ Sequence 4816 BP; 1275 A; 1285 C; 1265 G; 991 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 13; Length 4816;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTACTTCATAGTCTT 17
DB 811 TTACTTCATAGTCTT 796

RESULT 62
AEA61137
ID AEA61137 standard; DNA; 127567 BP.
XX AC AEA61137;
XX 25-AUG-2005 (first entry)
XX DE Human BCAS1 gene genomic sequence SEQ ID NO:47.
XX KW DNA methylation; biomarker; cancer; gene; ds; BCAS1.
XX OS Homo sapiens.
XX PN US2005130172-A1.
XX PD 16-JUN-2005.
XX 27-JAN-2004; 2004US-00765790.
XX PR 16-DEC-2003; 2003US-00737082.

PA	(FARB) BAYER CORP.
XX	
PI	Beard C, Burgess C, Gannon A, Harvey J, Lechner JF, Li Z;
XX	
DR	WPI: 2005-456991/46.
DR	GENBANK; NM_003657.
XX	
PT	Identifying nucleic acid sequences as biomarker for disease, by
PT	identifying nucleic acid sequences comprising methylated CpG site and
PT	down-regulated in diseased cells and comparing its expression level with
PT	demethylated nucleic acid.
XX	
PS	Claim 11; SEQ ID NO 47; 27pp; English.
XX	
CC	The invention relates to a method (M1) for identifying one or more
CC	nucleic acid sequences useful as a biomarker for a disease to be
CC	detected. (M1) involves identifying nucleic acid sequences comprising
CC	methylated CpG site in promoter-first exon region and that are down-
CC	regulated in diseased cells, comparing expression level of nucleic acid
CC	sequences with that of demethylated nucleic acid sequences and
CC	identifying nucleic acid sequences exhibiting increase in expression
CC	after demethylation. Also described: (1) detecting (M2) the presence or
CC	stage of a disease in a subject, which involves determining the degree of
CC	methylation of one or more CpG sites on nucleic acid sequences in a
CC	biological sample obtained from the subject, and determining the presence
CC	of, predisposition to, or stage of the disease in the subject based on
CC	the degree of methylation; (2) monitoring the onset, progression, or
CC	regression of a disease in a subject; (3) determining the efficacy of a
CC	test compound for inhibiting a disease in a subject; and (4) a kit (1)
CC	useful for diagnosis, prognosis, staging, monitoring, and therapeutic
CC	treatment of a disease. (M1) is useful for identifying one or more
CC	nucleic acid sequences useful as a biomarker for a disease to be
CC	detected, where the nucleic acid sequences are useful for detecting, the
CC	presence or stage of a disease such as cancer e.g. colorectal cancer in a
CC	subject. The present sequence represents a specifically claimed human
CC	genomic sequence for use in the method of the invention. Note - The
CC	sequence data for this patent is not represented in the printed
CC	specification but was obtained in electronic format from the USPTO web
CC	site.
XX	
SQ	Sequence 127567 BP; 36913 A; 25847 C; 25973 G; 38834 T; 0 U; 0 Other;
	Query Match 84.2%; Score 16; DB 14; Length 127567;
	Best Local Similarity 100.0%; Pred. No. 1.7e+03;
	Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps
Qy	3 TTACTTCATAGTCTTT 18
Db	110607 TTACTTCATAGTCTTT 110622
RESULT 63	
AAAC36870/c	
ID	AAAC36870 standard; DNA; 483 BP.
XX	
AC	AAAC36870;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 15361.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway; metabolic pathway;
KW	promoter; termination sequence; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	

DR WPI; 2004-053041/05.

Dr P-PSDB; ADK16127.
 XX New recombinant cancer-associated genes, such as KCMU9, useful for
 PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,
 PT cervical, or skin cancers, lymphomas, or leukemia.
 XX
 PS Claim 5; SEQ ID NO 78; 251pp; English.
 XX
 CC The invention comprises then amino acid and coding sequences of cancer-
 CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention
 CC also comprises the Nanoarchaeum equitans genome. The DNA and protein
 CC sequences of the invention are useful for diagnosing and treating cancer
 CC (e.g. carcinoma, lymphoma, or leukaemia). The present DNA sequence
 CC represents a Nanoarchaeum equitans CA gene of the invention.
 XX
 SQ Sequence 654 BP; 274 A; 87 C; 137 G; 156 T; 0 U; 0 Other;
 Query Match 83.2%; Score 15.8; DB 12; Length 654;
 Best Local Similarity 89.5%; Pred. No. 1.5e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGTCTTTG 19
 Db 84 CTTTAAATTGATAGTCTTTG 66
 RESULT 65
 ADN73034/C
 ID ADN73034 standard; cDNA; 1131 BP.
 XX
 AC ADN73034;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 929.
 XX
 KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
 KW growth regulator; animal feed product; thale cress;
 KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2004035798-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 20-OCT-2003; 2003WO-EP011658.
 XX
 PR 18-OCT-2002; 2002EP-00079408.
 XX
 PA (CROP-) CROPPDESIGN NV.
 XX
 PI Inze D, De Veylder L, Vlieghe K;
 XX
 DR WPI; 2004-348466/32.
 DR P-PSDB; ADN73035.
 XX
 XX Altering plant characteristics, useful for producing plants for enzyme or
 PT pharmaceutical production comprises modifying in a plant, expression of
 PT one or more nucleic acids and/or modifying level or activity of one or
 PT more proteins.
 XX
 PS Claim 1; SEQ ID NO 929; 134pp; English.
 XX
 CC This invention relates to a novel method for altering one or more plant
 CC characteristics. Specifically, it refers to identifying genes that are up
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
 CC alter plant characteristics accordingly. The present invention describes
 CC generating transgenic plants for the production of growth regulators,
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
 CC the altered plant characteristics are selected from increased yield or
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture

or physiology, altered endoreduplication, biochemistry, signal
 transduction, storage lipid mobilisation and/or altered photosynthesis,
 each relative to the corresponding wild type plants. Accordingly, these
 sequences can also be useful as positive or negative selectable markers
 during transformation of cells or tissues. The identified genes play a
 role in a variety of biological processes such as DNA replication, cell
 wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polynucleotide sequence is thale cress cDNA
 CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa
 CC transcription factor, given in an exemplification of the invention.
 XX
 SQ Sequence 1131 BP; 363 A; 183 C; 284 G; 301 T; 0 U; 0 Other;
 Query Match 83.2%; Score 15.8; DB 12; Length 1131;
 Best Local Similarity 89.5%; Pred. No. 1.6e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGTCTTTG 19
 Db 484 CTTCACTCCATAGTCTTTG 466
 RESULT 66
 AAS25932/C
 ID AAS25932 standard; cDNA; 1199 BP.
 XX
 AC AAS25932;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human cDNA encoding a novel secreted protein, Seq ID 111.
 XX
 KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
 KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
 KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;
 KW preservative; antiproliferative.
 XX
 OS Homo sapiens.
 XX
 PN WO200155322-A2.
 XX
 PD 02-AUG-2001.
 XX
 PR 17-JAN-2001; 2001WO-US001341.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.

PR	14-AUG-2000;	2000US-02232667P;
PR	14-AUG-2000;	2000US-02232688P;
PR	14-AUG-2000;	2000US-02232709P;
PR	14-AUG-2000;	2000US-02232747P;
PR	14-AUG-2000;	2000US-02232757P;
PR	14-AUG-2000;	2000US-02232758P;
PR	14-AUG-2000;	2000US-02232759P;
PR	18-AUG-2000;	2000US-02262799P;
PR	18-AUG-2000;	2000US-02262878P;
PR	22-AUG-2000;	2000US-02266818P;
PR	22-AUG-2000;	2000US-02266868P;
PR	22-AUG-2000;	2000US-02271828P;
PR	22-AUG-2000;	2000US-02272009P;
PR	30-AUG-2000;	2000US-02289248P;
PR	01-SEP-2000;	2000US-02282878P;
PR	01-SEP-2000;	2000US-02293438P;
PR	01-SEP-2000;	2000US-02293444P;
PR	01-SEP-2000;	2000US-02293445P;
PR	03-SEP-2000;	2000US-02295099P;
PR	05-SEP-2000;	2000US-02295133P;
PR	06-SEP-2000;	2000US-02304378P;
PR	06-SEP-2000;	2000US-02304388P;
PR	08-SEP-2000;	2000US-02312428P;
PR	08-SEP-2000;	2000US-02312438P;
PR	08-SEP-2000;	2000US-02312444P;
PR	08-SEP-2000;	2000US-02314138P;
PR	08-SEP-2000;	2000US-02314144P;
PR	08-SEP-2000;	2000US-02320808P;
PR	08-SEP-2000;	2000US-02320818P;
PR	12-SEP-2000;	2000US-02321968P;
PR	12-SEP-2000;	2000US-02323977P;
PR	14-SEP-2000;	2000US-02323988P;
PR	14-SEP-2000;	2000US-02323998P;
PR	14-SEP-2000;	2000US-02324000P;
PR	14-SEP-2000;	2000US-02324018P;
PR	14-SEP-2000;	2000US-02324063P;
PR	14-SEP-2000;	2000US-02330648P;
PR	14-SEP-2000;	2000US-02330658P;
PR	21-SEP-2000;	2000US-02342238P;
PR	21-SEP-2000;	2000US-02342748P;
PR	25-SEP-2000;	2000US-02345978P;
PR	25-SEP-2000;	2000US-02349988P;
PR	26-SEP-2000;	2000US-02354848P;
PR	27-SEP-2000;	2000US-02358348P;
PR	27-SEP-2000;	2000US-02358368P;
PR	29-SEP-2000;	2000US-02363278P;
PR	29-SEP-2000;	2000US-02363678P;
PR	29-SEP-2000;	2000US-02363688P;
PR	29-SEP-2000;	2000US-02363698P;
PR	29-SEP-2000;	2000US-02363708P;
PR	02-OCT-2000;	2000US-02368028P;
PR	02-OCT-2000;	2000US-02370378P;
PR	02-OCT-2000;	2000US-02370388P;
PR	02-OCT-2000;	2000US-02370398P;
PR	02-OCT-2000;	2000US-02370408P;
PR	13-OCT-2000;	2000US-02399358P;
PR	13-OCT-2000;	2000US-02399378P;
PR	20-OCT-2000;	2000US-02409568P;
PR	20-OCT-2000;	2000US-02412218P;
PR	20-OCT-2000;	2000US-02417858P;
PR	20-OCT-2000;	2000US-02417868P;
PR	20-OCT-2000;	2000US-02417878P;
PR	20-OCT-2000;	2000US-02418088P;
PR	20-OCT-2000;	2000US-02418098P;
PR	20-OCT-2000;	2000US-02418268P;
PR	01-NOV-2000;	2000US-02464617P;
PR	08-NOV-2000;	2000US-02464674P;
PR	08-NOV-2000;	2000US-02464675P;
PR	08-NOV-2000;	2000US-02464676P;
PR	08-NOV-2000;	2000US-02464677P;
PR	08-NOV-2000;	2000US-02464678P;
PR	08-NOV-2000;	2000US-02465238P;
PR	08-NOV-2000;	2000US-02465248P;
PR	08-NOV-2000;	2000US-02465258P;
PR	08-NOV-2000;	2000US-02465268P;

PR	08-NOV-2000;	2000US-0245526P.
PR	08-NOV-2000;	2000US-02445527P.
PR	08-NOV-2000;	2000US-02445528P.
PR	08-NOV-2000;	2000US-02445532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249279P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251588P.
PR	06-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-02551479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251389P.
PR	08-DEC-2000;	2000US-0251390P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259578P.

(HUMA-) HUMAN GENOME SCI. INC.

Rosen CA. Barash SC. Ruben SM:

WPI; 2001-488783/53.
P-PSDB: AAU15945.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 1: SEQ ID NO 111: 980pp: English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis the polypeptides can also be used

CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 83.2%; Score 15.8; DB 4; Length 1199;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
||||| ||||| |||||
Db 692 CTTTATTTCATAGCCTTTG 674

RESULT 67
ADB81647
ID ADB81647 standard; DNA; 1258 BP.

XX AC ADB81647;

XX DT 04-DEC-2003 (first entry)

XX OS Human ovarian specific gene SEQ ID NO:20.

XX DE ds; gene; human; ovarian specific gene; cytostatic; ovarian cancer;
XX KW ovary; vaccine; gene therapy; OSG; 5p25.3b.

XX OS Homo sapiens.

XX PN WO2003020897-A2.

XX PD 13-MAR-2003.

XX PF 29-AUG-2002; 2002WO-US027727.

XX PR 31-AUG-2001; 2001US-0316307P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Sun Y, Liu C, Salceda S;

XX WPI; 2003-290181/28.

DR P-PSDB; ADB81590.

XX New ovarian specific nucleic acid, useful as a vaccine, in gene therapy,
PT or for identifying, diagnosing, monitoring, staging, imaging and treating
PT ovarian cancer and non-cancerous disease states of the ovary.

XX Example 1; Page 162-163; 220pp; English.

XX The invention relates to a novel isolated ovarian specific nucleic acid.
CC A nucleic acid sequence of the invention has cytostatic activity. The
CC ovarian specific nucleic acid or a polypeptide encoded by it is useful
CC for identifying, diagnosing, monitoring, staging, imaging and treating
CC ovarian cancer and non-cancerous disease states of the ovary. The ovarian
CC specific nucleic acid molecule or polypeptide is also useful as a
CC vaccine, in gene therapy, for producing transgenic animals and cells, or
CC producing engineered ovarian tissue for treatment and research. The
CC antibody that specifically binds to the polypeptide is useful for
CC treating a patient with ovarian cancer, particularly for inducing an
CC immune response against the ovarian cancer cell expressing the nucleic
CC acid molecule or polypeptide. The sequences shown in ADB81628-ADB81688
CC represent human ovarian specific genes (OSG's) of the invention.

XX SQ Sequence 1258 BP; 460 A; 200 C; 215 G; 382 T; 0 U; 1 Other;

Query Match 83.2%; Score 15.8; DB 10; Length 1258;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
||||| ||||| |||||

Db 1232 CTTTATTTCATATCTTTG 1250

RESULT 68

ID ABX73273/c
ABX73273 standard; DNA; 1259 BP.

XX AC ABX73273;

XX DT 18-MAR-2003 (first entry)

XX OS Human novel polynucleotide #101.

XX DE Human; gene; ds; neural disorder; immune system disorder; renal disorder;
XX KW muscular disorder; respiratory disease; reproductive disorder;
XX KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX KW hyperproliferative disorder; inflammatory disease; allergic reaction;
XX KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX KW haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.

XX PN US2002132753-A1.

XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 22-AUG-2000; 2000US-0226868P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139898P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140334P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

```
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 23-OCT-1999; 99US-0161405P.
PR 23-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      83.2%  Score 15.6; DB 3; Length 1394;
Best Local Similarity 89.5%  Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 552 CTTCACTCCATAGTCTTTG 534

RESULT 71
AAC33429/c
ID AAC33429 standard; DNA; 1398 BP.
XX
AC AAC33429;
XX
DT 17-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3005.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134376P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
```

PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146336P.
PR 02-AUG-1999; 99US-0146338P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160788P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 83.2%; Score 15.8; DB 3; Length 1398;

Best Local Similarity 89.5%; Pred. No. 1.6e+03;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGTCTTTG 19
||| ||| ||| ||| ||| |||
Db 555 CTTCACTCCATAGTCTTTG 537

RESULT 72

AAI60916/c

ID AAI60916 standard; cDNA; 1401 BP.

XX AC AAI60916;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 4905.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW Peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

XX OS Homo sapiens.

XX XX WO200153312-A1.

XX XX 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PSDB; AAM41760.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.

XX Claim 1; SEQ ID NO 4905; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1401 BP; 473 A; 286 C; 320 G; 322 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 4; Length 1401;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTG 19
||||| ||||| |||||
Db 829 CTTTATTCATAGCCTTG 811

RESULT 73
ADB83252/c
ID ADB83252 standard; cDNA; 1599 BP.
XX
AC ADB83252;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA sequence useful for the treatment of cancer (SeqID 1465).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.

PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Ckvenjakov R, Dickson M, Drmanac S, Labat I, Leehkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
XX New polynucleotides derived from human prostate, useful for modulating
XX immune response to prevent or treat cancer.
XX
XX Claim 1; SEQ ID NO 1465; 189pp; English.

XX This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/ or vaccines they can be used to modulate the immune response for the

CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTE: These sequences are not given in
CC the specification but are provided on the WIPO website.

XX
SQ Sequence 1599 BP; 528 A; 327 C; 378 G; 366 T; 0 U; 0 Other;
Query Match 83.2%; Score 15.8; DB 9; Length 1599;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTG 19
||||| ||||| |||||
Db 456 CTTTATTCATAGCCTTG 438

RESULT 74
AAS64753/c
ID AAS64753 standard; cDNA; 1734 BP.
XX
AC AAS64753;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #557.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR P-PSDB; ABG00566.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 557; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1734 BP; 507 A; 419 C; 458 G; 350 T; 0 U; 0 Other;
Best Local Similarity 83.2%; Score 15.8; DB 5; Length 1734;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGCTTTG 19
||||| ||||| |||||
Db 1272 CTTTATTTCATAGCCTTTG 1254

RESULT 75
ADN38895/c
ID ADN38895 standard; cDNA; 1758 BP.

XX AC ADN38895;

XX DT 17-JUN-2004 (first entry)

XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:213.

XX KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO2003042661-A2.

XX PD 22-MAY-2003.

XX PF 13-NOV-2002; 2002WO-US036810.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-035250P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-0368809P.

XX PR 04-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-039775P.

XX PR 22-JUL-2002; 2002US-0397845P.

XX PR 09-SEP-2002; 2002US-0409450P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX PR WPI; 2003-468649/44.

XX PR P-PSDB; ADN38896.

XX PT Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.

XX PS Claim 8; SEQ ID NO 213; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38893-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX SQ Sequence 1758 BP; 554 A; 372 C; 412 G; 420 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 11; Length 1758;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGCTTTG 19
||||| ||||| |||||
Db 615 CTTTATTTCATAGCCTTTG 597

RESULT 76
ACC43802/c
ID ACC43802 standard; DNA; 1868 BP.

XX AC ACC43802;

XX DT 11-AUG-2003 (first entry)

XX DE Nucleotide sequence of plant regulatory polynucleotide.

XX KW Plant; regulatory polynucleotide; endothelium; albumen; seed; seed testa;
XX flavonoid; tannin; cellulose; maize; rape seed; colour; nutrition; ss.

XX OS Arabidopsis thaliana.

XX PN FR2828210-A1.

XX PD 07-FEB-2003.

XX PR 01-AUG-2001; 2001FR-00010365.

XX PR 01-AUG-2001; 2001FR-00010365.

XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX (CNRS) CNRS CENT NAT RECH SCI.

XX PI Lepintec L, Debeaujon I, Devic M, Caboche M;

XX DR WPI; 2003-335321/32.

XX PT New regulatory nucleic acid, useful for preparing transgenic plants with
XX altered seed testa, provides specific expression of attached genes in
XX seed endothelium and albumen.

XX PS Disclosure; Page 80; 85pp; French.

XX The present sequence represents a plant regulatory polynucleotide, which
CC was used to construct plasmid pBan2. The present sequence is a fragment
CC of the sequence given in ACC43789, which comprises regulatory sequences
CC that allow expression of a polynucleotide specifically in the endothelium
CC and/or albumen of plant seeds. The regulatory polynucleotide, and derived
CC expression cassettes or vectors, are used to produce transgenic plants in
CC which development of seed testa is altered, especially for reducing the

CC size or number of seeds, particularly complete absence of mature, fertile
CC seeds, or modulating contents of flavonoids, tannins and cellulose
CC (especially to produce maize and rape seeds with less intense yellow
CC colour, or to alter nutritional properties)

XX Sequence 1868 BP; 592 A; 304 C; 402 G; 570 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 8; Length 1868;

Best Local Similarity 89.5%; Pred. No. 1.6e+03;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19

|||||
Db 532 CTTCACTCCATAGTCTTTG 514

RESULT 77

ADJ41386/C

ID ADJ41386 standard; cDNA; 2000 BP.

XX AC ADJ41386;

DT 06-MAY-2004 (first entry)

XX DE Plant cDNA #2386.

XX Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.

XX OS Eukaryota.

XX PN US2004016025-A1.

XX PD 23-JAN-2004.

XX PF 26-SEP-2002; 2002US-00260238.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 26-SEP-2001; 2001US-0325448P.

XX PR 04-APR-2002; 2002US-0370620P.

XX (BUDW/) BUDWORTH P.

PA (MOUG/) MOUGHAMER T.

PA (BRIG/) BRIGGS S P.

PA (COOP/) COOPER B.

PA (GLAZ/) GLAZEBROOK J.

PA (GOFF/) GOFF S A.

PA (KATA/) KATAGIRI F.

PA (KREP/) KREPS J.

PA (PROV/) PROVART N.

PA (RICK/) RICHE D.

PA (ZHUT/) ZHU T.

XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;

PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;

XX WPI; 2004-190374/18.

XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.

XX Claim 1; SEQ ID NO 2386; 230pp; English.

XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant

CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2000 BP; 542 A; 542 C; 499 G; 417 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 12; Length 2000;

Best Local Similarity 89.5%; Pred. No. 1.7e+03;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19

|||||
Db 675 CTTTGCTTCATAGTCTTCG 657

RESULT 78

ABL60217/C

ID ABL60217 standard; cDNA; 2021 BP.

XX AC ABL60217;

XX DT 29-JUL-2002 (first entry)

XX DE Human encoding zinc finger protein 59 cDNA SEQ ID NO 1.

XX KW Human; zinc finger protein 59; disease; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 191..1789

XX FT /*tag= a

XX FT /product= "zinc finger protein 59"

XX CN1293203-A.

XX PD 02-MAY-2001.

XX PF 18-OCT-1999; 99CN-00116988.

XX PR 18-OCT-1999; 99CN-00116988.

XX PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX PI Mao Y, Xie Y;

XX WPI; 2001-418901/45.

XX P-PSDB; ABB77494.

XX Polypeptide-human zinc finger protein 59 and polynucleotide for coding
PT this polypeptide.

XX Claim 6; Page 16-17 (Disclosure); 25pp; Chinese.

XX The invention relates to human zinc finger protein 59 (ABB77494), the
CC polynucleotide (ABL60217) encoding it, the process for preparing it, its
CC application in treating several diseases, the antagonist and medical
CC action and coding the new applications of the polynucleotide are
CC disclosed

XX SQ Sequence 2021 BP; 680 A; 398 C; 478 G; 465 T; 0 U; 0 Other;

```
Query Match      83.2%; Score 15.8; DB 4; Length 2021;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| |||||
Db 646 CTTTATTTCATAGCCTTTG 628

RESULT 79
AAH16484/c
ID AAH16484 standard; cDNA; 2063 BP.
XX
AC AAH16484;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:15507.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
FN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
XX
PR 11-JAN-2000; 2000JP-00118776.
XX
PR 02-MAY-2000; 2000JP-00183767.
XX
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 15507; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 2063 BP; 691 A; 406 C; 489 G; 477 T; 0 U; 0 Other;

Query Match      83.2%; Score 15.8; DB 4; Length 2021;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| |||||
Db 646 CTTTATTTCATAGCCTTTG 628

RESULT 80
ABL20545
ID ABL20545 standard; DNA; 2094 BP.
XX
AC ABL20545;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 13108.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 13108; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2094 BP; 478 A; 569 C; 546 G; 501 T; 0 U; 0 Other;

Query Match      83.2%; Score 15.8; DB 4; Length 2094;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| |||||
Db 784 CTCTACTTCATAGTCTTTG 802

RESULT 81
ABL05491
ID ABL05491 standard; cDNA; 2165 BP.
XX
XX
AC ABL05491;
```

```
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10955.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX OS pharmaceutical; gene; ss.
XX PN Drosophila melanogaster.
XX PD WO200171042-A2.
XX PF 27-SEP-2001.
XX PI 23-MAR-2001; 2001WO-US009231.
XX PT 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB61388.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX Claim 1; SEQ ID NO 10955; 2lpp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pt_sequences
XX
XX SQ Sequence 2165 BP; 501 A; 590 C; 556 G; 518 T; 0 U; 0 Other;
Query Match 83.2%; Score 15.8; DB 4; Length 2165;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CTTTACTTCATAGCTTTTG 19
Db 855 CTCTACTTCATGCTTTTG 873
RESULT 82
ABQ72580/c
ID ABQ72580 standard; cDNA; 2175 BP.
AC ABQ72580;
XX 03-SEP-2002 (first entry)
XX Human MDDT encoding cDNA SEQ ID NO 132.
XX Human; MDDT; disease detection and treatment molecule polynucleotide;
XX proliferative disorder; hepatitis; psoriasis; cancer; AIDS;
XX autoimmune disorder; inflammatory disorder; allergy; multiple sclerosis;
XX rheumatoid arthritis; transgenic; gene therapy; antiarteriosclerotic;
XX hepatotropic; antiinflammatory; antipsoriatic; cytostatic; anti-HIV;
XX antiallergic; antianaemic; antiasthmatic; antiatherosclerotic; antitout;
XX neuroprotective; antirheumatic; antiarthritic; gene; ss.
XX Homo sapiens.
```

```
XX WO200240715-A2.
XX 23-MAY-2002.
XX 06-SEP-2001; 2001WO-US027628.
XX 05-SEP-2000; 2000US-0229747P.
XX 05-SEP-2000; 2000US-0229748P.
XX 05-SEP-2000; 2000US-0229749P.
XX 05-SEP-2000; 2000US-0229750P.
XX 05-SEP-2000; 2000US-0229751P.
XX 05-SEP-2000; 2000US-0230583P.
XX 06-SEP-2000; 2000US-0230505P.
XX 06-SEP-2000; 2000US-0230514P.
XX 06-SEP-2000; 2000US-0230515P.
XX 06-SEP-2000; 2000US-0230517P.
XX 06-SEP-2000; 2000US-0230518P.
XX 06-SEP-2000; 2000US-0230519P.
XX 06-SEP-2000; 2000US-0230595P.
XX 06-SEP-2000; 2000US-0230597P.
XX 06-SEP-2000; 2000US-0230598P.
XX 06-SEP-2000; 2000US-0230599P.
XX 06-SEP-2000; 2000US-0230610P.
XX 06-SEP-2000; 2000US-0230865P.
XX 06-SEP-2000; 2000US-0230988P.
XX 07-SEP-2000; 2000US-0230989P.
XX 07-SEP-2000; 2000US-0231163P.
XX 07-SEP-2000; 2000US-0231167P.
XX (INCY-) INCYTE GENOMICS INC.
XX Jackson S, Lincoln SE, Altus CM, Dufour GE, Chalup MS;
XX Hillman JL, Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE;
XX Dahl CR, Momiyama MG, Bradley DL, Rohatgi SD, Harris B;
XX Roseberry AM, Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V;
XX Daffo A, Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI; 2002-527544/56.
XX P-PSDB; ABP51363.
XX Novel human disease detection and treatment polypeptide, useful in
XX diagnosis, prevention or treatment of cell proliferative disorders e.g.
XX arteriosclerosis, cirrhosis and an autoimmune/inflammatory disorder e.g.
XX AIDS.
XX Claim 1; Page 381; 618pp; English.
XX The invention relates to an isolated human disease detection and
XX treatment (MDDT) polypeptide (I) selected from a polypeptide having a
XX sequence selected from 254 sequences (ABP51231-ABP51484) given in the
XX specification, a naturally occurring polypeptide comprising a sequence
XX having at least 90% identity to (I) or a biologically active or
XX immunogenic fragment of (I). (I) is useful for screening a compound for
XX effectiveness as an agonist or antagonist, for screening a compound that
XX specifically binds (I) or modulates the activity of (I), and for
XX preparing a polyclonal or monoclonal antibody by hybridoma technology.
XX Nucleic acids (II) (ABQ72449-ABQ72700) encoding (I) are useful for
XX screening a compound for effectiveness in altering expression of a target
XX polynucleotide comprising. Oligonucleotides and antibodies are useful for
XX detecting MDDT in a sample or for assessing toxicity of a test compound,
XX in a diagnostic test for a condition or a disease associated with the
XX expression of MDDT in a biological sample, for detecting (I) in a sample,
XX and for purifying (I) from a sample. A composition comprising (I), an
XX agonist or antagonist is useful for treating a disease or condition
XX associated with decreased or increased expression of functional MDDT. (I)
XX or (II) are useful for diagnosing, treating or preventing disorders
XX associated with aberrant expression of MDDT, where the disorders are
XX selected from a cell proliferative disorder such as arteriosclerosis,
XX cirrhosis, hepatitis, psoriasis, and cancer and an
XX autoimmune/inflammatory disorder such as AIDS, Addison's disease,
XX allergy, anaemia, asthma, atherosclerosis, gout, multiple sclerosis or
```

CC rheumatoid arthritis. (II) are useful for creating knockin humanised
 CC animals or transgenic animals to model human diseases, in somatic or
 CC germline gene therapy, to generate a transcript image of a tissue or cell
 CC type, for detecting differences in the chromosomal location due to
 CC translocation or inversion among normal, carrier or affected individuals
 CC and as hybridisation probes for mapping naturally occurring genomic
 CC sequences

SQ Sequence 2175 BP; 726 A; 415 C; 526 G; 508 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 6; Length 2175;
 Best Local Similarity 89.5%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
 ||||| ||||| ||||| |||||
 Db 677 CTTTACTTCATAGCTTTG 659

RESULT 83

ID AD226902/c
 ID AD226902 standard; DNA; 2208 BP.

XX AC AD226902;

DT 30-JUN-2005 (first entry)

DE Adeno-associated virus DNA SEQ ID NO 52.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

OS Adeno-associated virus.

XX WO2005033321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004WO-US028817.

XX PR 30-SEP-2003; 2003US-0508226P.

XX PR 29-APR-2004; 2004US-0566546P.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX DR WPI; 2005-285437/29.

XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.

XX Claim 46; SEQ ID NO 52; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents an adeno-associated virus DNA.

SQ Sequence 2208 BP; 599 A; 629 C; 526 G; 454 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 14; Length 2208;
 Best Local Similarity 89.5%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
 ||||| ||||| ||||| |||||
 Db 1596 CTTTCTTCATCGCTTTG 1578

RESULT 84

ID AD226918/c
 ID AD226918 standard; DNA; 2208 BP.

XX AC AD226918;

DT 30-JUN-2005 (first entry)

DE Adeno-associated virus DNA SEQ ID NO 68.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

OS Adeno-associated virus.

XX WO2005033321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004WO-US028817.

XX PR 30-SEP-2003; 2003US-0508226P.

XX PR 29-APR-2004; 2004US-0566546P.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX DR WPI; 2005-285437/29.

XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.

XX Claim 15; SEQ ID NO 68; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents an adeno-associated virus DNA.

SQ Sequence 2208 BP; 592 A; 641 C; 544 G; 431 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 14; Length 2208;
 Best Local Similarity 89.5%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19
 ||||| ||||| ||||| |||||
 Db 1596 CTTTCTTCATCGCTTTG 1578

RESULT 85

```

AD226920/c
ID AD226920 standard; DNA; 2208 BP.
AC AD226920;
XX
XX 30-JUN-2005 (first entry)
DT
XX
XX Adeno-associated virus DNA SEQ ID NO 70.
DE
XX
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
XX WO2005033321-A2.
XX
XX 14-APR-2005.
XX
XX 30-SEP-2004; 2004WO-US028817.
XX
XX 30-SEP-2003; 2003US-0508226P.
XX
XX 29-APR-2004; 2004US-0566546P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
XX Claim 50; SEQ ID NO 57; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX
XX Sequence 2208 BP; 587 A; 640 C; 548 G; 433 T; 0 U; 0 Other;
XX
XX Query Match 83.2%; Score 15.8; DB 14; Length 2208;
XX Best Local Similarity 89.5%; Pred. No. 1.7e+03;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CTTTACTTCATAGTCCTTTG 19
XX 1596 CTTTCTTCATCGTCCTTTG 1578
XX
XX
XX RESULT 86
XX AD226907/c
XX ID AD226907 standard; DNA; 2208 BP.
XX
XX AC AD226907;
XX
XX 30-JUN-2005 (first entry)
XX
XX Adeno-associated virus DNA SEQ ID NO 57.
XX
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW

```

```

KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
XX WO2005033321-A2.
XX
XX 14-APR-2005.
XX
XX 30-SEP-2004; 2004WO-US028817.
XX
XX 30-SEP-2003; 2003US-0508226P.
XX
XX 29-APR-2004; 2004US-0566546P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
XX Claim 50; SEQ ID NO 57; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX
XX Sequence 2208 BP; 601 A; 626 C; 527 G; 454 T; 0 U; 0 Other;
XX
XX Query Match 83.2%; Score 15.8; DB 14; Length 2208;
XX Best Local Similarity 89.5%; Pred. No. 1.7e+03;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CTTTACTTCATAGTCCTTTG 19
XX 1596 CTTTCTTCATCGTCCTTTG 1578
XX
XX
XX RESULT 87
XX AD226903/c
XX ID AD226903 standard; DNA; 2208 BP.
XX
XX AC AD226903;
XX
XX 30-JUN-2005 (first entry)
XX
XX Adeno-associated virus DNA SEQ ID NO 53.
XX
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
XX WO2005033321-A2.
XX

```

```

Pd 14-APR-2005.
XX 30-SEP-2004; 2004WO-US028817.
XX 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX WPI; 2005-285437/29.
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX WPI; 2005-285437/29.
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX Disclosure; SEQ ID NO 53; 569pp; English.
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2208 BP; 602 A; 628 C; 525 G; 453 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 14; Length 2208;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATGCTTTG 19
Db 1596 CTTTCTTCATGCTTTG 1578

RESULT 89
ADZ26901/C
ID ADZ26901 standard; DNA; 2208 BP.
XX AC ADZ26901;
XX 30-JUN-2005 (first entry)
XX Adeno-associated virus DNA SEQ ID NO 51.
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX antiarthritic; neuroprotective; antinflammatory; antidiabetic;
XX antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX Adeno-associated virus.
XX OS
XX WO2005033321-A2.
XX 14-APR-2005.
XX 30-SEP-2004; 2004WO-US028817.
XX 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX WPI; 2005-285437/29.
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX Claim 50; SEQ ID NO 55; 569pp; English.

```


CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2208 BP; 601 A; 626 C; 526 G; 455 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 14; Length 2208;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1596 CTTTCTTCATCGTCTTTG 1578

RESULT 90
ADZ26906/c
ID ADZ26906 standard; DNA; 2208 BP.

XX AC ADZ26906;

DT 30-JUN-2005 (first entry)

XX DE Adeno-associated virus DNA SEQ ID NO 56.

XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX KW antiarthritic; neuroprotective; antinflammatory; antidiabetic;
XX KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX OS Adeno-associated virus.

XX PN WO2005033321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004WO-US028817.

XX PR 30-SEP-2003; 2003US-0508226P.

XX PR 29-APR-2004; 2004US-0566546P.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX WPI; 2005-285437/29.

XX CC New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.

XX PS Claim 50; SEQ ID NO 56; 569pp; English.

XX CC The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.

XX SQ Sequence 2208 BP; 602 A; 628 C; 527 G; 451 T; 0 U; 0 Other;
Query Match 83.2%; Score 15.8; DB 14; Length 2208;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1596 CTTTCTTCATCGTCTTTG 1578

RESULT 91

ADZ26900/c

ID ADZ26900 standard; DNA; 2208 BP.

XX AC ADZ26900;

DT 30-JUN-2005 (first entry)

XX DE Adeno-associated virus DNA SEQ ID NO 50.

XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX KW antiarthritic; neuroprotective; antinflammatory; antidiabetic;
XX KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX OS Adeno-associated virus.

XX PN WO2005033321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004WO-US028817.

XX PR 30-SEP-2003; 2003US-0508226P.

XX PR 29-APR-2004; 2004US-0566546P.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX WPI; 2005-285437/29.

XX CC New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.

XX PS Claim 50; SEQ ID NO 50; 569pp; English.

XX CC The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.

XX SQ Sequence 2208 BP; 601 A; 627 C; 524 G; 456 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 14; Length 2208;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1596 CTTTCTTCATCGTCTTTG 1578

```

RESULT 92
AD226908/c
ID AD226908 standard; DNA; 2208 BP.
XX AC
XX AD226908;
XX DT
XX 30-JUN-2005 (first entry)
XX DE
XX Adeno-associated virus DNA SEQ ID NO 58.
XX KW
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX antiarthritic; neuroprotective; antinflammatory; antidiabetic;
XX antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX OS
XX Adeno-associated virus.
XX PN
XX WO2005033321-A2.
XX PD
XX 14-APR-2005.
XX PF
XX 30-SEP-2004; 2004WO-US028817.
XX PR
XX 30-SEP-2003; 2003US-0508226P.
XX PR
XX 29-APR-2004; 2004US-0566546P.
XX PA
XX (UYPE-) UNIV PENNSYLVANIA.
XX PI
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX DR
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX PS
XX Claim 7; SEQ ID NO 58; 569pp; English.
XX CC
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX SQ
XX Sequence 2208 BP; 602 A; 628 C; 525 G; 453 T; 0 U; 0 Other;
XX Query Match 83.2%; Score 15.8; DB 14; Length 2208;
XX Best Local Similarity 89.5%; Pred. No. 1.7e+03;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY
XX 1 CTTTACTTCATGCTTTG 19
XX DB
XX 1596 CTTTCTTCATGCTTTG 1578
XX RESULT 93
AD226909/c
ID AD226909 standard; DNA; 2208 BP.
XX AC
XX AD226909;
XX DT
XX 30-JUN-2005 (first entry)
XX DE
XX Adeno-associated virus DNA SEQ ID NO 54.
XX KW
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX antiarthritic; neuroprotective; antinflammatory; antidiabetic;
XX antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX OS
XX Adeno-associated virus.
XX PN
XX WO2005033321-A2.
XX PD
XX 14-APR-2005.
XX PF
XX 30-SEP-2004; 2004WO-US028817.
XX PR
XX 30-SEP-2003; 2003US-0508226P.
XX PR
XX 29-APR-2004; 2004US-0566546P.
XX PA
XX (UYPE-) UNIV PENNSYLVANIA.
XX PI
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX DR
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX PS
XX Claim 7; SEQ ID NO 58; 569pp; English.
XX CC
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX SQ
XX Sequence 2208 BP; 602 A; 628 C; 525 G; 453 T; 0 U; 0 Other;
XX Query Match 83.2%; Score 15.8; DB 14; Length 2208;
XX Best Local Similarity 89.5%; Pred. No. 1.7e+03;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY
XX 1 CTTTACTTCATGCTTTG 19
XX DB
XX 1596 CTTTCTTCATGCTTTG 1578
XX RESULT 94
AD226904/c
ID AD226904 standard; DNA; 2208 BP.
XX AC
XX AD226904;
XX DT
XX 30-JUN-2005 (first entry)
XX DE
XX Adeno-associated virus DNA SEQ ID NO 59.
XX KW
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX antiarthritic; neuroprotective; antinflammatory; antidiabetic;
XX antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX OS
XX Adeno-associated virus.
XX PN
XX WO2005033321-A2.
XX PD
XX 14-APR-2005.
XX PF
XX 30-SEP-2004; 2004WO-US028817.
XX PR
XX 30-SEP-2003; 2003US-0508226P.
XX PR
XX 29-APR-2004; 2004US-0566546P.
XX PA
XX (UYPE-) UNIV PENNSYLVANIA.
XX PI
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX DR
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX PS
XX Claim 7; SEQ ID NO 59; 569pp; English.
XX CC
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX SQ
XX Sequence 2208 BP; 602 A; 627 C; 526 G; 453 T; 0 U; 0 Other;
XX Query Match 83.2%; Score 15.8; DB 14; Length 2208;
XX Best Local Similarity 89.5%; Pred. No. 1.7e+03;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX QY
XX 1 CTTTACTTCATGCTTTG 19
XX DB
XX 1596 CTTTCTTCATGCTTTG 1578

```

```
XX PN WO2005033321-A2.
XX PD 14-APR-2005.
XX PF 30-SEP-2004; 2004WO-US028817.
XX PR 30-SEP-2003; 2003US-0508226P.
XX PR 29-APR-2004; 2004US-0566546P.
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX DR WPI; 2005-285437/29.
XX PT New adeno-associated virus (AAV) clade comprising at least three AAV
XX PT members, useful for preventing and/or treating arthritis, multiple
XX PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX PT infection and cancer.
XX PS Disclosure; SEQ ID NO 54; 569pp; English.
XX CC The invention relates to an adeno-associated virus (AAV) clade comprising
XX CC at least three AAV members, where each member of the AAV clade is
XX CC phylogenetically related as determined using a Neighbor-Joining heuristic
XX CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX CC correction distance measurement of no more than 0.05. The methods and
XX CC compositions of the present invention are useful for the prevention
XX CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX CC colitis. The present sequence represents an adeno-associated virus DNA.
XX SQ Sequence 2208 BP; 601 A; 626 C; 527 G; 454 T; 0 U; 0 Other;
Query Match 83.2%; Score 15.8; DB 14; Length 2208;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATGCTCTTG 19
Db ||||| ||||| ||||| |||||
1596 CTTTCTTCATGCTCTTG 1578

RESULT 95
ADZ26919/C
ID ADZ26919 standard; DNA; 2208 BP.
XX AC ADZ26919;
XX DT 30-JUN-2005 (first entry)
XX DE Adeno-associated virus DNA SEQ ID NO 69.
XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
XX KW antiporiatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.
XX OS Adeno-associated virus.
XX PN WO2005033321-A2.
XX PD 14-APR-2005.
XX PF 30-SEP-2004; 2004WO-US028817.
XX PR 30-SEP-2003; 2003US-0508226P.
XX PR 29-APR-2004; 2004US-0566546P.
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX DR WPI; 2005-285437/29.
XX PT New adeno-associated virus (AAV) clade comprising at least three AAV
XX PT members, useful for preventing and/or treating arthritis, multiple
XX PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX PT infection and cancer.
XX PS Disclosure; SEQ ID NO 54; 569pp; English.
XX CC The invention relates to an adeno-associated virus (AAV) clade comprising
XX CC at least three AAV members, where each member of the AAV clade is
XX CC phylogenetically related as determined using a Neighbor-Joining heuristic
XX CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX CC correction distance measurement of no more than 0.05. The methods and
XX CC compositions of the present invention are useful for the prevention
XX CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX CC colitis. The present sequence represents an adeno-associated virus DNA.
XX SQ Sequence 2208 BP; 601 A; 626 C; 527 G; 454 T; 0 U; 0 Other;
Query Match 83.2%; Score 15.8; DB 14; Length 2208;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATGCTCTTG 19
Db ||||| ||||| ||||| |||||
1596 CTTTCTTCATGCTCTTG 1578

RESULT 96
AAH72833/C
ID AAH72833 standard; cDNA; 2366 BP.
XX AC AAH72833;
XX DT 19-SEP-2001 (first entry)
XX DE Human cervical cancer marker nucleic acid 4107.
XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX OS Homo sapiens.
XX PN WO200142467-A2.
XX PD 14-JUN-2001.
XX PF 08-DEC-2000; 2000WO-US033312.
XX PR 08-DEC-1999; 99US-0169681P.
XX PR 21-DEC-1999; 99US-0171350P.
XX PR 14-MAR-2000; 2000US-0189315P.
XX PR 12-MAY-2000; 2000US-0203791P.
XX PR 09-JUN-2000; 2000US-0210600P.
XX PR 21-JUL-2000; 2000US-0220114P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Deeds J, Berger A, Zhao X;
XX DR WPI; 2001-375006/39.
XX PT New isolated nucleic acid for diagnosing and treating cervical cancer and
XX PT for assessing and detecting compounds for treating the cancer.
XX PS Claim 1; Page 877-878; 1051pp; English.
XX
```

CC The invention relates to novel genes (AAH68727-AAH73383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful: to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy

XX SQ Sequence 2366 BP; 746 A; 476 C; 578 G; 559 T; 0 U; 7 Other;

Query Match 83.2%; Score 15.8; DB 4; Length 2366;

Best Local Similarity 89.5%; Pred. No. 1.7e+03;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

||||| ||||| ||||| ||||| |||||

695 CTTTATTTCATAGCCTTTG 677

RESULT 97

ACC43789/C

ID ACC43789 standard; DNA; 2376 BP.

XX AC ACC43789;

XX 11-AUG-2003 (first entry)

XX Nucleotide sequence of plant regulatory polynucleotide.

XX Plant; regulatory polynucleotide; endothelium; albumen; seed; seed testa; flavonoid; tannin; cellulose; maize; rape seed; colour; nutrition; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

FT Promoter 2172..2312

FT /*tag= b

FT CAAT_signal 2172..2180

FT /*tag= a

FT TATA_signal 2306..2312

FT /*tag= c

XX FR2828210-A1.

XX 07-FEB-2003.

XX 01-AUG-2001; 2001FR-00010365.

XX 01-AUG-2001; 2001FR-00010365.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Lepiniec L, Debeaujon I, Devic M, Caboche M;

XX WPI; 2003-335321/32.

XX New regulatory nucleic acid, useful for preparing transgenic plants with altered seed testa, provides specific expression of attached genes in seed endothelium and albumen.

XX Claim 1; Page 75-76; 85pp; French.

XX The present sequence represents a plant regulatory polynucleotide, comprising regulatory sequences that allow expression of a polynucleotide specifically in the endothelium and/or albumen of plant seeds. The present sequence was used to produce plasmid pBan1. The regulatory polynucleotide, and derived expression cassettes or vectors, are used to produce transgenic plants in which development of seed testa is altered, especially for reducing the size or number of seeds, particularly complete absence of mature, fertile seeds, or modulating contents of flavonoids, tannins and cellulose (especially to produce maize and rape

CC seeds with less intense yellow colour, or to alter nutritional properties)

XX SQ Sequence 2376 BP; 804 A; 382 C; 463 G; 727 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 8; Length 2376;

Best Local Similarity 89.5%; Pred. No. 1.7e+03;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

||||| ||||| ||||| ||||| |||||

Db 1040 CTTCACTCCATAGCTTTG 1022

RESULT 98

AAI59130/C

ID AAI59130 standard; cDNA; 2412 BP.

XX AC AAI59130;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 1333.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PSDB; AAM39974.

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

XX Claim 1; SEQ ID NO 1333; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX SQ Sequence 2412 BP; 793 A; 476 C; 570 G; 573 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 4; Length 2412;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| |||||
Db 746 CTTTATTTTCATAGCCTTTG 728

RESULT 99
ADQ99353/c
ID ADQ99353 standard; cDNA; 2412 BP.

XX AC ADQ99353;

XX DT 23-SEP-2004 (first entry)

XX DE DNA encoding human GPCR-like protein seqid 1023.

XX KW ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
XX KW antidiabetic; GPCR-like protein; ophthalmic disorder;
XX KW neurological disorder; immunological disorder; nephritic disorder;
XX KW hormonal dysfunction; cancer; atherosclerosis; diabetes;
XX KW molecular weight marker; food supplement; human; ss.

XX OS Homo sapiens.

XX FN US659662-B1.

XX PD 27-MAY-2003.

XX PF 19-JUL-2000; 2000US-00620312.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Drmanac RT;

XX DR WPI; 2001-442255/47.

XX PT New G-protein-coupled receptor-like polypeptides and polynucleotides,
XX PT useful for treating diseases of ophthalmic, neurological, immunological
XX PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
XX PT and diabetes.

XX PS Example 2; SEQ ID NO 1023; 92pp; English.

XX CC The invention describes an isolated polynucleotide (I) comprising a fully
XX CC defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
XX CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
XX CC given in the specification, its translated or protein coding portion, its
XX CC extracellular portion or its active domain. The GPCR-like polypeptides
XX CC and polynucleotides are useful for the treatment of diseases of
XX CC ophthalmic, neurological, immunological and nephritic systems. They may
XX CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
XX CC diabetes. The antibodies are useful for detecting or quantitating the
XX CC polypeptide in tissue. The polypeptides can also be used as molecular
XX CC weight markers and as a food supplement. This sequence represents a human
XX CC polynucleotide of the invention.

XX SQ Sequence 2412 BP; 793 A; 476 C; 570 G; 573 T; 0 U; 0 Other;

'Query Match 83.2%; Score 15.8; DB 5; Length 2412;

Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| |||||
Db 746 CTTTATTTTCATAGCCTTTG 728

RESULT 100
ADB49113/c
ID ADB49113 standard; cDNA; 2412 BP.

XX AC ADB49113;

XX DT 04-DEC-2003 (first entry)

XX DE Novel human cDNA SEQ ID NO 1023.

XX KW ss; cancer; neurodegenerative disease; human.

XX OS Homo sapiens.

XX PN US2003104529-A1.

XX PD 05-JUN-2003.

XX PF 04-JAN-2002; 2002US-00037270.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 19-JUL-2000; 2000US-00620312.

XX PA (ZHOU/) ZHOU P.

XX PA (TANG/) TANG Y T.

XX PA (LIUC/) LIU C.

XX PA (ASUN/) ASUNDI V.

XX PA (DRMA/) DRMANAC R T.

XX PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;

XX WPI; 2003-678194/64.

XX PT New polynucleotide, useful for treating diseases e.g., cancer or
XX PT neurodegenerative diseases.

XX PS Claim 1; SEQ ID NO 1023; 99pp; English.

XX CC The invention relates to a polynucleotide comprising a sequence given in
XX CC the specification, or its mature protein-coding portion, or its
XX CC complement. The polynucleotide is useful for treating diseases e.g.,
XX CC cancer or neurodegenerative diseases and many others listed in the
XX CC specification. The present sequence represents a novel human cDNA. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from USPTO
XX CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.

XX SQ Sequence 2412 BP; 793 A; 476 C; 570 G; 573 T; 0 U; 0 Other;

Query Match 83.2%; Score 15.8; DB 9; Length 2412;
Best Local Similarity 89.5%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| |||||
Db 746 CTTTATTTTCATAGCCTTTG 728

Search completed: June 10, 2006, 15:32:51
Job time : 172.483 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:19:28 ; Search time 2344.6 Seconds
(without alignments)
453.155 Million cell updates/sec

Title: US-10-600-816-32

Perfect score: 19

Sequence: 1 ctttacttcattagctttg 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est3:*

3: gb_est4:*

4: gb_est5:*

5: gb_est6:*

6: gb_hic:*

7: gb_est2:*

8: gb_est7:*

9: gb_est8:*

10: gb_est9:*

11: gb_gse1:*

12: gb_gse2:*

13: gb_gse3:*

14: gb_gse4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	100.0	158	2	BI063311 IL3-UT011
C 2	19	100.0	307	7	BE188833 CM2-BN030
C 3	19	100.0	341	7	BE188817 CM2-BN030
C 4	19	100.0	365	7	BE188831 CM2-BN030
C 5	19	100.0	367	1	AA112374 zB68e10.1
C 6	19	100.0	374	7	BE003240 QV4-BN009
C 7	19	100.0	384	7	BE188835 CM2-BN030
C 8	19	100.0	396	7	BE188828 CM2-BN030
C 9	19	100.0	404	7	BE188840 CM2-BN030
C 10	19	100.0	424	7	BE188805 CM2-BN030
C 11	19	100.0	449	7	BE188807 CM2-BN030
C 12	19	100.0	453	5	CD672930 fgl8a05.Y
C 13	19	100.0	458	7	BE188871 CM2-BN030
C 14	19	100.0	466	7	BE188859 CM2-BN030
C 15	19	100.0	467	7	BE188852 CM2-BN030
C 16	19	100.0	571	3	BP253378 BP253378
C 17	19	100.0	577	4	CB152636 K-EST0209
C 18	19	100.0	580	3	BP265234 BP265234
C 19	19	100.0	580	3	BP268151 BP268151

582	3	BP256144	BP256144
582	3	BP363758	BP363758
584	3	BP256147	BP256147
584	3	BP256147	BP256147
621	4	CB130950	CB130950
621	4	CB130950	CB130950
630	7	BE747107	BE747107
630	7	BE747107	BE747107
637	7	BM843051	BM843051
637	7	BM843051	BM843051
659	7	BF568108	BF568108
659	7	BF568108	BF568108
702	2	BG251131	BG251131
702	2	BG251131	BG251131
851	3	BQ887352	BQ887352
851	3	BQ887352	BQ887352
855	4	CA454746	CA454746
855	4	CA454746	CA454746
877	2	BM011054	BM011054
877	2	BM011054	BM011054
899	3	BQ958341	BQ958341
899	3	BQ958341	BQ958341
901	2	BG831564	BG831564
901	2	BG831564	BG831564
918	4	BX372160	BX372160
918	4	BX372160	BX372160
923	3	BU543952	BU543952
923	3	BU543952	BU543952
940	3	BQ685729	BQ685729
940	3	BQ685729	BQ685729
962	4	BX370558	BX370558
962	4	BX370558	BX370558
1023	3	BQ067433	BQ067433
1023	3	BQ067433	BQ067433
1830	6	CR597125	CR597125
1830	6	CR597125	CR597125
490	14	DU976902	DU976902
490	14	DU976902	DU976902
811	13	DU069153	DU069153
811	13	DU069153	DU069153
916	5	CF780868	CF780868
916	5	CF780868	CF780868
1125	2	BI193620	BI193620
1125	2	BI193620	BI193620
400	14	DE194703	DE194703
400	14	DE194703	DE194703
429	11	AQ435685	AQ435685
429	11	AQ435685	AQ435685
458	7	AW664308	AW664308
458	7	AW664308	AW664308
495	8	CN050801	CN050801
495	8	CN050801	CN050801
529	2	BG009082	BG009082
529	2	BG009082	BG009082
565	10	DT490210	DT490210
565	10	DT490210	DT490210
580	3	BP267627	BP267627
580	3	BP267627	BP267627
603	1	AM058428	AM058428
603	1	AM058428	AM058428
630	11	AZ353695	AZ353695
630	11	AZ353695	AZ353695
635	14	AG158060	AG158060
635	14	AG158060	AG158060
644	11	AZ840748	AZ840748
644	11	AZ840748	AZ840748
657	10	DV148009	DV148009
657	10	DV148009	DV148009
703	10	DT482123	DT482123
703	10	DT482123	DT482123
718	3	BU290329	BU290329
718	3	BU290329	BU290329
735	9	DN485086	DN485086
735	9	DN485086	DN485086
740	10	DV140694	DV140694
740	10	DV140694	DV140694
765	8	CO081056	CO081056
765	8	CO081056	CO081056
766	8	CV276848	CV276848
766	8	CV276848	CV276848
779	10	DT478689	DT478689
779	10	DT478689	DT478689
779	10	DT483764	DT483764
779	10	DT483764	DT483764
789	10	DT483588	DT483588
789	10	DT483588	DT483588
802	8	CV258213	CV258213
802	8	CV258213	CV258213
830	12	CG329092	CG329092
830	12	CG329092	CG329092
834	8	CV242157	CV242157
834	8	CV242157	CV242157
837	10	DT506297	DT506297
837	10	DT506297	DT506297
839	10	DT487539	DT487539
839	10	DT487539	DT487539
844	10	DT487258	DT487258
844	10	DT487258	DT487258
852	8	CV254849	CV254849
852	8	CV254849	CV254849
859	10	DT489385	DT489385
859	10	DT489385	DT489385
877	8	CV245752	CV245752
877	8	CV245752	CV245752
877	10	DT522272	DT522272
877	10	DT522272	DT522272
887	10	DT521741	DT521741
887	10	DT521741	DT521741
897	12	CG329079	CG329079
897	12	CG329079	CG329079
912	7	BE548141	BE548141
912	7	BE548141	BE548141
955	12	CG147825	CG147825
955	12	CG147825	CG147825
976	12	CG330802	CG330802
976	12	CG330802	CG330802
459	4	CB828264	CB828264
459	4	CB828264	CB828264
470	11	AQ413031	AQ413031
470	11	AQ413031	AQ413031
552	4	CB827270	CB827270
552	4	CB827270	CB827270
578	11	AZ483135	AZ483135
578	11	AZ483135	AZ483135
811	14	CT259201	CT259201
811	14	CT259201	CT259201
820	2	BI819265	BI819265
820	2	BI819265	BI819265
884	14	DU846802	DU846802
884	14	DU846802	DU846802
894	14	BU970629	BU970629
894	14	BU970629	BU970629
907	3	BU931252	BU931252
907	3	BU931252	BU931252
2921	6	AK171788	AK171788
2921	6	AK171788	AK171788
2923	6	AK172413	AK172413
2923	6	AK172413	AK172413
2942	6	AK170599	AK170599
2942	6	AK170599	AK170599
3785	6	AK170599	AK170599
3785	6	AK170599	AK170599
3796	6	AK076230	AK076230
3796	6	AK076230	AK076230

C 93	16.4	86.3	91	14	DE242715	DE242715 Trifolium	166	16.4	86.3	579	2	BI808034	BI808034 B009D01 O
C 94	16.4	86.3	172	11	BH497760	BH497760 BOHPC14TF	167	16.4	86.3	580	9	CX728508	CX728508 OSJNEF08B
C 95	16.4	86.3	205	7	BB419641	BB419641 BB419641	168	16.4	86.3	583	14	DX225245	DX225245 OR_Aba010
C 96	16.4	86.3	245	7	BB171187	BB171187 BB171187	C 169	16.4	86.3	584	4	CB469287	CB469287 sn05_E08.
C 97	16.4	86.3	279	7	BB176170	BB176170 BB176170	C 170	16.4	86.3	587	7	AW632065	AW632065 91594 MAR
C 98	16.4	86.3	290	7	BB027420	BB027420 BB027420	C 171	16.4	86.3	586	11	BZ160278	BZ160278 CH230-484
C 99	16.4	86.3	293	13	CZ443147	CZ443147 CHMD-GT.2	C 172	16.4	86.3	596	14	DE242238	DE242238 Trifolium
C 100	16.4	86.3	300	9	DN761057	DN761057 G.hir-5te	C 173	16.4	86.3	598	1	AT003383	AT003383 AT003383
C 101	16.4	86.3	315	9	DN917229	DN917229 MCF7RNAL2	C 174	16.4	86.3	599	5	C1146461	C1146461 C1146461
C 102	16.4	86.3	353	8	CN772565	CN772565 tad83c07.	C 175	16.4	86.3	600	5	C1142926	C1142926 C1142926
C 103	16.4	86.3	359	7	BE653744	BE653744 UI-M-AH1-	C 176	16.4	86.3	603	5	CF963268	CF963268 8489rsice
C 104	16.4	86.3	360	2	BF851901	BF851901 RC0-EN008	C 177	16.4	86.3	605	4	CB681611	CB681611 OSJNEF08B
C 105	16.4	86.3	374	7	AW619310	AW619310 678 MARC	C 178	16.4	86.3	607	5	CF957669	CF957669 2506rsice
C 106	16.4	86.3	375	8	CO986449	CO986449 UNC-pd10e	C 179	16.4	86.3	614	10	DR383629	DR383629 13814038
C 107	16.4	86.3	376	5	C1295663	C1295663 C1295663	C 180	16.4	86.3	616	10	DY381632	DY381632 ZO_Eg000
C 108	16.4	86.3	377	3	BP822638	BP822638 BP822638	C 181	16.4	86.3	621	13	CL338739	CL338739 RPCI44_26
C 109	16.4	86.3	377	7	AV812536	AV812536 AV812536	C 182	16.4	86.3	623	12	CE227329	CE227329 tigr-gss-
C 110	16.4	86.3	377	14	DE035053	DE035053 Branchios	C 183	16.4	86.3	628	1	AU173558	AU173558 AU173558
C 111	16.4	86.3	379	3	BP819921	BP819921 BP819921	C 184	16.4	86.3	633	3	EW283904	EW283904 BW283904
C 112	16.4	86.3	380	1	AJ275279	AJ275279 AJ275279	C 185	16.4	86.3	634	3	EW258766	EW258766 BW258766
C 113	16.4	86.3	380	3	BP841431	BP841431 BP841431	C 186	16.4	86.3	635	5	CF990387	CF990387 20566rsic
C 114	16.4	86.3	383	5	CJ272433	CJ272433 CJ272433	C 187	16.4	86.3	639	5	C1312345	C1312345 C1312345
C 115	16.4	86.3	387	3	BP638004	BP638004 BP638004	C 188	16.4	86.3	641	4	CA043515	CA043515 ssalmgfo0
C 116	16.4	86.3	388	7	AV804530	AV804530 AV804530	C 189	16.4	86.3	645	4	CB676269	CB676269 OSJNE12G
C 117	16.4	86.3	390	8	CO940306	CO940306 UNC-pd45f	C 190	16.4	86.3	646	11	AZ908443	AZ908443 RPCI_24-2
C 118	16.4	86.3	393	8	CO943468	CO943468 UNC-p2mm4	C 191	16.4	86.3	647	14	CR344614	CR344614 mcel-78C2
C 119	16.4	86.3	409	3	BP650514	BP650514 BP650514	C 192	16.4	86.3	649	14	DE200349	DE200349 Branchios
C 120	16.4	86.3	416	3	BP648131	BP648131 BP648131	C 193	16.4	86.3	650	3	BW371856	BW371856 BW371856
C 121	16.4	86.3	425	4	CB696439	CB696439 AMGNUC-S	C 194	16.4	86.3	650	12	BZ980125	BZ980125 PUGK89TB
C 122	16.4	86.3	432	1	AA83518	AA83518 n4400C.8	C 195	16.4	86.3	651	8	CN627664	CN627664 taeg7a10.
C 123	16.4	86.3	437	7	AA813264	AA813264 AA813264	C 196	16.4	86.3	656	7	BB626886	BB626886 BB626886
C 124	16.4	86.3	441	4	CB000663	CB000663 S345L_E06	C 197	16.4	86.3	658	13	CL417560	CL417560 ZMMBB042
C 125	16.4	86.3	441	7	BB749797	BB749797 BB749797	C 198	16.4	86.3	661	14	DE008136	DE008136 Branchios
C 126	16.4	86.3	445	1	AJ673572	AJ673572 AJ673572	C 199	16.4	86.3	662	13	CG638853	CG638853 OA_Aba017
C 127	16.4	86.3	445	1	AJ673572	AJ673572 AJ673572	C 200	16.4	86.3	662	13	DE279279	DE279279 Oryzias 1
C 128	16.4	86.3	445	2	BI378101	BI378101 BFLG3_001	C 201	16.4	86.3	672	11	DE094838	DE094838 Oryzias 1
C 129	16.4	86.3	445	4	CA967782	CA967782 CGL03a12	C 202	16.4	86.3	673	11	BH840108	BH840108 LMC20001
C 130	16.4	86.3	446	14	DU715599	DU715599 pyn80c01.	C 203	16.4	86.3	679	2	BI381080	BI381080 BFLG1_002
C 131	16.4	86.3	449	1	AJ606484	AJ606484 AJ606484	C 204	16.4	86.3	680	10	DY377298	DY377298 ZO_Eg000
C 132	16.4	86.3	450	3	BQ484865	BQ484865 pmp1c.pk0	C 205	16.4	86.3	682	7	BB026216	BB026216 BB026216
C 133	16.4	86.3	450	5	CI130197	CI130197 CI130197	C 206	16.4	86.3	682	14	DE210480	DE210480 Branchios
C 134	16.4	86.3	450	5	CI167554	CI167554 CI167554	C 207	16.4	86.3	682	10	DY377297	DY377297 ZO_Eg000
C 135	16.4	86.3	453	5	CI330753	CI330753 CI330753	C 208	16.4	86.3	689	5	CX958229	CX958229 4058525 B
C 136	16.4	86.3	464	5	CI333466	CI333466 CI333466	C 209	16.4	86.3	692	9	CX576888	CX576888 TT8000310
C 137	16.4	86.3	467	5	CI330042	CI330042 CI330042	C 210	16.4	86.3	693	3	BW211091	BW211091 BW211091
C 138	16.4	86.3	467	12	CE267483	CE267483 tigr-gss-	C 211	16.4	86.3	699	5	CF536506	CF536506 UI-M-G10-
C 139	16.4	86.3	474	1	AJ697151	AJ697151 AJ697151	C 212	16.4	86.3	699	11	AZ818667	AZ818667 2M0088P13
C 140	16.4	86.3	483	1	AI659635	AI659635 t06d03.x	C 213	16.4	86.3	700	3	BU230488	BU230488 603946248
C 141	16.4	86.3	483	1	AJ696739	AJ696739 AJ696739	C 214	16.4	86.3	700	14	CR492024	CR492024 mch2-165P
C 142	16.4	86.3	493	2	BI337559	BI337559 361024 MA	C 215	16.4	86.3	702	8	CR286747	CR286747 CR286747
C 143	16.4	86.3	498	5	CF962401	CF962401 14228rsic	C 216	16.4	86.3	705	14	AG317889	AG317889 Mus muscu
C 144	16.4	86.3	500	2	BM303517	BM303517 SNEST4a61	C 217	16.4	86.3	708	11	AZ870628	AZ870628 2M0183K12
C 145	16.4	86.3	500	5	CI153666	CI153666 CI153666	C 218	16.4	86.3	717	12	CE509868	CE509868 tigr-gss-
C 146	16.4	86.3	500	5	CI282578	CI282578 CI282578	C 219	16.4	86.3	718	4	CB676268	CB676268 OSJNE12G
C 147	16.4	86.3	500	5	CI283923	CI283923 CI283923	C 220	16.4	86.3	719	4	CB633504	CB633504 OSJNE12G
C 148	16.4	86.3	500	5	CI284185	CI284185 CI284185	C 221	16.4	86.3	719	13	DU261314	DU261314 109857423
C 149	16.4	86.3	517	3	BQ490719	BQ490719 EST00018	C 222	16.4	86.3	722	3	BU306798	BU306798 603735211
C 150	16.4	86.3	517	12	CG469306	CG469306 ZMMBB026	C 223	16.4	86.3	722	13	AG576801	AG576801 Mus muscu
C 151	16.4	86.3	526	5	CI059357	CI059357 CI059357	C 224	16.4	86.3	723	13	CW866360	CW866360 shs2kd41-
C 152	16.4	86.3	530	10	TI51106	TI51106 crs750 lamb	C 225	16.4	86.3	725	4	CB680289	CB680289 OSJNEF04M
C 153	16.4	86.3	532	11	AZ018260	AZ018260 RPCI-23-3	C 226	16.4	86.3	728	8	CO045480	CO045480 UI-M-H00-
C 154	16.4	86.3	542	5	CI154127	CI154127 CI154127	C 227	16.4	86.3	729	11	BZ219219	BZ219219 CH230-250
C 155	16.4	86.3	543	12	CE703898	CE703898 tigr-gss-	C 228	16.4	86.3	731	4	CB629279	CB629279 OSJNEB05G
C 156	16.4	86.3	546	11	BH124978	BH124978 RPCI-24-3	C 229	16.4	86.3	731	14	AG316438	AG316438 Mus muscu
C 157	16.4	86.3	550	5	CI264439	CI264439 CI264439	C 230	16.4	86.3	733	11	BZ441150	BZ441150 BONQV85TF
C 158	16.4	86.3	551	5	CI061376	CI061376 CI061376	C 231	16.4	86.3	734	3	BW266934	BW266934 BW266934
C 159	16.4	86.3	558	3	BU547594	BU547594 GM880013A	C 232	16.4	86.3	734	10	DY381631	DY381631 ZO_Eg000
C 160	16.4	86.3	559	5	CF163109	CF163109 B0721A04-	C 233	16.4	86.3	735	8	CX039922	CX039922 REC0085
C 161	16.4	86.3	560	4	CA998795	CA998795 S345M_B12	C 234	16.4	86.3	736	9	CX581753	CX581753 TT8000313
C 162	16.4	86.3	567	3	BW208977	BW208977 BW208977	C 235	16.4	86.3	737	4	CB417623	CB417623 590347 MA
C 163	16.4	86.3	569	1	AV383195	AV383195 AV383195	C 236	16.4	86.3	739	8	CR286780	CR286780 CR286780
C 164	16.4	86.3	570	11	AQ482795	AQ482795 RPCI-11-2	C 237	16.4	86.3	740	11	BH604106	BH604106 BOHKS60TR
C 165	16.4	86.3	571	5	CI286683	CI286683 CI286683	C 238	16.4	86.3	748	4	CB965002	CB965002 NL14_H08

c 239	16.4	86.3	749	4	CB629278	CB629278 OSIIEB05G	312	16	84.2	363	7	AW089200	AW089200 xc82h03.x
c 240	16.4	86.3	749	14	AG608488	Mus muscu	c 313	16	84.2	377	2	BF902761	BF902761 IL5-MT020
c 241	16.4	86.3	750	10	DY355381	ZO_Ec000	314	16	84.2	388	10	R95594	R95594 vt83b03.r1
c 242	16.4	86.3	753	4	BX901285	BX901285	315	16	84.2	394	2	BF931285	BF931285 IL5-NT022
c 243	16.4	86.3	761	1	AJ939284	AJ939284	316	16	84.2	397	11	AZ895933	AZ895933 RPC1-24-2
c 244	16.4	86.3	761	5	CJ360903	CJ360903	317	16	84.2	399	3	BQ355757	BQ355757 MR4-HT105
c 245	16.4	86.3	763	4	CB665471	OSJNEd110	318	16	84.2	402	2	BF929975	BF929975 IL2-NT019
c 246	16.4	86.3	763	10	DY373477	ZO_EF000	c 319	16	84.2	417	2	BI009081	BI009081 RC5-RT005
c 247	16.4	86.3	768	11	AZ585844	1M0391N17	320	16	84.2	423	2	BI040545	BI040545 PM2-NT023
c 248	16.4	86.3	773	4	CA344958	CA344958 675460 NC	c 321	16	84.2	434	2	BI006224	BI006224 RC5-RT005
c 249	16.4	86.3	776	8	CX115507	CX115507 EI099P01	c 322	16	84.2	434	2	BI055233	BI055233 PM2-GN049
c 250	16.4	86.3	781	12	BZ877744	CH240_293	323	16	84.2	445	2	BI040755	BI040755 PM2-MT023
c 251	16.4	86.3	783	12	CC385229	CC385229 PUMHE39TD	324	16	84.2	479	1	AA731088	AA731088 nw68a08.s
c 252	16.4	86.3	785	4	CB683217	OSJNEf11F	c 325	16	84.2	485	1	AA767230	AA767230 nz80h01.s
c 253	16.4	86.3	787	13	DU330350	DU330350 109826835	326	16	84.2	487	4	CB140358	CB140358 K-EST0193
c 254	16.4	86.3	790	14	SP6105F5	CT009012 Ernus cab	c 327	16	84.2	487	9	DA788764	DA788764 DA788764
c 255	16.4	86.3	792	10	DV905297	DV905297 L80213.CR	c 328	16	84.2	498	1	AI423271	AI423271 tf41b12.x
c 256	16.4	86.3	794	11	BZ507381	BZ507381 BOOAP25TR	c 329	16	84.2	503	9	DB218795	DB218795 DB218795
c 257	16.4	86.3	795	14	CNS06VXT	AL417735 T7 end of	c 330	16	84.2	513	7	AW504148	AW504148 UI-HF-BN0
c 258	16.4	86.3	796	1	AM036979	AM036979 AM036979	c 331	16	84.2	515	2	BG998827	BG998827 MR4-HT105
c 259	16.4	86.3	800	12	BZ991832	PUGGS89TD	c 332	16	84.2	517	3	BQ331458	BQ331458 MR4-HT105
c 260	16.4	86.3	802	4	CB680288	OSJNEf04M	c 333	16	84.2	517	3	BQ331458	BQ331458 MR4-HT105
c 261	16.4	86.3	802	5	CK770025	CK770025 958142 MA	c 334	16	84.2	517	9	DA739575	DA739575 DA739575
c 262	16.4	86.3	805	14	CT400359	Sus scrof	c 335	16	84.2	522	2	BI007597	BI007597 RCS-RT005
c 263	16.4	86.3	806	14	DE246518	Trifolium	c 336	16	84.2	526	9	DB207723	DB207723 DB207723
c 264	16.4	86.3	809	13	CZ679409	OM_Ba023	c 337	16	84.2	527	2	BG998725	BG998725 MR4-HT105
c 265	16.4	86.3	811	4	CB683630	OSJNEf12E	c 338	16	84.2	532	2	BG994981	BG994981 MR4-HT105
c 266	16.4	86.3	820	11	BH540150	BH540150 BOGRF70TF	c 339	16	84.2	532	2	EG995058	EG995058 MR4-HT105
c 267	16.4	86.3	823	8	CO081126	CO081126 GR_Ea44P	c 340	16	84.2	532	3	BM836789	BM836789 K-EST0112
c 268	16.4	86.3	826	11	BH499073	BH499073 BOHQ840TF	c 341	16	84.2	540	9	DA121074	DA121074 DA121074
c 269	16.4	86.3	832	14	CT433195	CT433195 Sus scrof	c 342	16	84.2	542	2	DA240210	DA240210 MR2-HT116
c 270	16.4	86.3	837	12	CG339532	CG339532 OGWL225TV	c 343	16	84.2	546	9	DA240210	DA240210
c 271	16.4	86.3	842	1	AM013574	AM013574 603284056	c 344	16	84.2	552	9	DB216814	DB216814
c 272	16.4	86.3	846	2	BI655241	BI655241 603284056	c 345	16	84.2	557	9	DA124895	DA124895
c 273	16.4	86.3	846	3	BU277995	BU277995 603863722	c 346	16	84.2	560	9	DA788853	DA788853 DA788853
c 274	16.4	86.3	849	9	DN525334	DN525334 1269945 M	c 347	16	84.2	561	1	AI017576	AI017576 OU23f12.x
c 275	16.4	86.3	854	10	DY377439	ZO_Eg000	c 348	16	84.2	562	9	DA804780	DA804780 DA804780
c 276	16.4	86.3	860	14	CR048957	CR048957 Forward s	c 349	16	84.2	564	9	DA803663	DA803663 DA803663
c 277	16.4	86.3	869	8	CR281936	CR281936	c 350	16	84.2	568	2	BF817683	BF817683 MR2-CI012
c 278	16.4	86.3	871	7	BF119250	BF119250 601757258	c 351	16	84.2	569	3	BQ355765	BQ355765 MR4-HT105
c 279	16.4	86.3	879	8	CR281882	CR281882	c 352	16	84.2	569	3	DA999045	DA999045 DA999045
c 280	16.4	86.3	887	12	CG926602	CG926602 MBELX11TR	c 353	16	84.2	572	9	DA145874	DA145874 DA145874
c 281	16.4	86.3	897	14	AG552005	Mus muscu	c 354	16	84.2	578	9	DA776955	DA776955 DA776955
c 282	16.4	86.3	901	14	CNS06V79	AL416779 T3 end of	c 355	16	84.2	580	9	DA246621	DA246621 DA246621
c 283	16.4	86.3	902	14	CR259817	CR259817 Forward s	c 356	16	84.2	589	4	DA415509	DA415509 DA415509
c 284	16.4	86.3	907	8	CO435899	CO435899 OSNR533_5	c 357	16	84.2	592	4	CA949914	CA949914 ir84b10.y
c 285	16.4	86.3	911	2	BF975685	BF975685 602246453	c 358	16	84.2	593	9	DB160401	DB160401 DB160401
c 286	16.4	86.3	912	14	CT297424	Sus scrof	c 359	16	84.2	597	9	DB163085	DB163085 DB163085
c 287	16.4	86.3	913	13	CL415133	CL415133 ZMMBB042	c 360	16	84.2	598	3	DB161419	DB161419 DB161419
c 288	16.4	86.3	928	12	CG253920	CG253920 OGYBT78TH	c 361	16	84.2	615	3	BQ638557	BQ638557 hd24h10.y
c 289	16.4	86.3	928	13	DU324650	DU324650 109844825	c 362	16	84.2	629	13	DU421247	DU421247 109842110
c 290	16.4	86.3	929	4	CB182892	CB182892 AGENCOURT	c 363	16	84.2	636	14	CT078518	CT078518 Sus scrof
c 291	16.4	86.3	937	13	CZ953316	CZ953316 269907 TO	c 364	16	84.2	644	7	BE217979	BE217979 hv32e07.x
c 292	16.4	86.3	1026	8	CX106353	CX106353 B1126D09	c 365	16	84.2	644	11	AZ630157	AZ630157 1M048JN03
c 293	16.4	86.3	1270	14	AG361013	AG361013 Mus muscu	c 366	16	84.2	647	8	BN043663	BN043663 v11_p48.d
c 294	16.4	86.3	1347	10	DV781684	DV781684 Hw Fat_50	c 367	16	84.2	648	2	BG896763	BG896763 HOA49-1-H
c 295	16.4	86.3	1496	6	AK017132	AK017132 Mus muscu	c 368	16	84.2	655	14	AG111592	AG111592 Pan trogl
c 296	16.4	86.3	2173	6	AK030561	AK030561	c 369	16	84.2	656	8	CN349289	CN349289 170005326
c 297	16.4	86.3	2508	6	CR749422	CR749422 Homo sapi	c 370	16	84.2	657	8	CV821937	CV821937 81095_130
c 298	16.4	86.3	3487	6	AK034784	AK034784 Mus muscu	c 371	16	84.2	674	2	BG576360	BG576360 602597079
c 299	16.4	86.3	3991	6	AK032751	AK032751 Mus muscu	c 372	16	84.2	686	8	CV824722	CV824722 84325_130
c 300	16	84.2	210	3	BQ431316	BQ431316 UI-HF-BN0	c 373	16	84.2	687	8	CN349292	CN349292 170005326
c 301	16	84.2	230	3	BQ354764	BQ354764 MR1-HT105	c 374	16	84.2	691	12	CE434109	CE434109 tigr-g88-
c 302	16	84.2	245	2	BF924924	BF924924 IL2-NT019	c 375	16	84.2	706	2	EG754245	EG754245 602707996
c 303	16	84.2	245	2	BF926061	BF926061 IL2-NT019	c 376	16	84.2	714	12	CE500834	CE500834 tigr-g88-
c 304	16	84.2	262	2	BG729682	BG729682 f082f04.y	c 377	16	84.2	740	2	BG819834	BG819834 602781890
c 305	16	84.2	301	7	AW277070	AW277070 xp61f01.x	c 378	16	84.2	750	13	CC320938	CC320938 TAM32-24K
c 306	16	84.2	302	13	CM078610	CM078610 104_395_1	c 379	16	84.2	764	13	CZ561843	CZ561843 l_p_bw015
c 307	16	84.2	305	2	BF912438	BF912438 IL2-UT007	c 380	16	84.2	792	8	CN349290	CN349290 170005998
c 308	16	84.2	305	2	BF912534	BF912534	c 381	16	84.2	795	3	BQ775293	BQ775293 UI-H-PH0-
c 309	16	84.2	328	3	BQ355775	BQ355775 MR4-HT105	c 382	16	84.2	797	8	CN349297	CN349297 170005316
c 310	16	84.2	360	2	BI007009	BI007009 RCS-RT005	c 383	16	84.2	806	3	BU748400	BU748400 CH3019_G
c 311	16	84.2	361	2	BI007003	BI007003 RCS-RT005	c 384	16	84.2	817	13	CZ988842	CZ988842 204534 TO

385 16 84.2 825 11 BZ165671 CH230-277
C 386 16 84.2 846 3 BU166308 AGENCOURT
387 16 84.2 853 11 BZ101394 CH230-146
C 388 16 84.2 867 13 DU467106 109844827
C 389 16 84.2 868 8 CX166444 HES2_40
C 390 16 84.2 868 13 CX2809822 OC_Ea018
C 391 16 84.2 870 3 BU521163 AGENCOURT
C 392 16 84.2 872 11 AQ740139 HS_5504_A
C 393 16 84.2 874 2 B1254161 602974803
C 394 16 84.2 879 5 CD516819 AGENCOURT
C 395 16 84.2 899 3 BO949071 AGENCOURT
C 396 16 84.2 918 1 AA767233 H280J05_8
C 397 16 84.2 919 13 DU46751 151607 To
C 398 16 84.2 926 3 BU181551 AGENCOURT
C 399 16 84.2 936 13 CZ253368 LQYCG73TV
C 400 16 84.2 948 2 BG823561 602728946
C 401 16 84.2 957 1 AL544435 AL544435
C 402 16 84.2 962 13 CZ990002 206995 To
C 403 16 84.2 1019 2 BG751278 602730219
C 404 16 84.2 1029 3 BQ050514 AGENCOURT
C 405 16 84.2 1070 12 CC239452 CH261-81F
C 406 16 84.2 1073 7 BE781343 601467064
C 407 16 84.2 2348 6 CR860695 Pongo Pvg
C 408 16 84.2 3273 14 DQ053676 Homo_Sapi
C 409 15.8 83.2 126 1 AA750976 ISKP0130
C 410 15.8 83.2 133 3 BQ971267 QHB6G11.Y
C 411 15.8 83.2 133 3 BU016839 QHE14C15.
C 412 15.8 83.2 133 3 BU025723 QHG11A24.
C 413 15.8 83.2 133 3 BU028999 QH6P24.Y
C 414 15.8 83.2 134 3 BU024345 QHF2605.Y
C 415 15.8 83.2 143 3 BU020514 QHE27K06.
C 416 15.8 83.2 144 3 BU016488 QHE13D10.
C 417 15.8 83.2 144 3 BU028874 QHHS124.Y
C 418 15.8 83.2 145 3 BQ090693 QHA11P17.
C 419 15.8 83.2 145 3 BQ091527 QHE10H18.
C 420 15.8 83.2 151 3 BQ976740 QH123J10.
C 421 15.8 83.2 153 3 BU023300 QHF10G14.
C 422 15.8 83.2 154 3 BU016161 QHE12E18.
C 423 15.8 83.2 171 3 BQ337820 PM3-NN108
C 424 15.8 83.2 194 11 BH50176 SALK_0371
C 425 15.8 83.2 203 7 BB000370 BB000370
C 426 15.8 83.2 205 6 CNO0821X Single re
C 427 15.8 83.2 219 7 BB585248 BB585248
C 428 15.8 83.2 221 7 BB007591 BB007591
C 429 15.8 83.2 230 11 BZ318180 BZ318180 hx40906.g
C 430 15.8 83.2 236 13 DU130833 DU130833 00591-C07
C 431 15.8 83.2 242 3 BU018701 QHE19G19.
C 432 15.8 83.2 263 13 CZ655648 OM_Ba020
C 433 15.8 83.2 294 11 AQ452170 HS_5092_A
C 434 15.8 83.2 306 12 CG093303 FULC06TD
C 435 15.8 83.2 312 1 AA558513 n171e04.s
C 436 15.8 83.2 312 6 CNO095YN BX043803 Single re
C 437 15.8 83.2 318 11 AQ929438 RPCI-23-2
C 438 15.8 83.2 311 5 CF107262 Shultzomi
C 439 15.8 83.2 333 5 CK635130 UI-M-HN0
C 440 15.8 83.2 339 4 CB685852 OSUNE16K
C 441 15.8 83.2 348 8 CX175123 D06_69-29
C 442 15.8 83.2 352 1 AA343440 EST49548
C 443 15.8 83.2 355 5 CF079138 QHK5F04.Y
C 444 15.8 83.2 361 9 CX947624 DH0A02ZE
C 445 15.8 83.2 362 2 BG184704 RST3525_A
C 446 15.8 83.2 362 11 AZ447482 AZ447482 IM0244J07
C 447 15.8 83.2 362 12 BZ765727 SALK_1338
C 448 15.8 83.2 364 10 DY183625 000930BEM
C 449 15.8 83.2 368 7 BF094233 CM1-UT004
C 450 15.8 83.2 368 13 CZ087448 OM_Ba008
C 451 15.8 83.2 369 8 CO310664 EK222236.
C 452 15.8 83.2 376 13 CB064276 OM_Ba006
C 453 15.8 83.2 378 4 BW927971 BW927971
C 454 15.8 83.2 378 11 AQ722301 HS_5210_B
C 455 15.8 83.2 381 5 CJ075839 CJ075839
C 456 15.8 83.2 384 2 BF830501 RC3-HT047
C 457 15.8 83.2 385 11 BH388352 AG-ND-143

387 8 CO950837 UMC-pd12c
389 5 CU081858 CU081858
393 11 AA820156 AA820156 L023628.5
393 11 AQ136283 HS_3056_B
394 8 CV266706 WS0203.B2
395 12 CG413973 ZMMBB028
396 12 CE497544 tigr-g88-
396 12 CE497544 tigr-g88-
403 3 BQ917019 BQ917019 QHB20A10.
404 3 BQ041147 GQ34C10.Y
405 10 W53093 W53093 mdi4b02.r1
406 5 CJ076945 CJ076945
411 4 BY503311 BY503311
415 4 AA774858 AA774858 a883d05.8
419 10 Z26091 Z26091 ATTS1363.Gr
430 11 AQ056080 AQ056080 CIT-HSP-2
430 11 B30856 HS-1103-A2-
430 13 CL605555 CH240_182
435 1 AA674959 vG51h09.r
435 1 AI015245 OY55A01.8
442 13 CZ019471 CH240_506
442 11 AZ505232 IM0345M16
445 11 AA163328 mt65D01.r
451 4 BY060450 BY060450
451 10 W59654 W59654 md73e06.r1
455 3 CF096892 QHN2H13.Y
455 5 CW364815 f8db001f0
461 11 A2290155 A2290155 RPCI-23-5
462 1 AI459221 tk1a11.x
467 7 AW235529 dn19c07.x
469 9 DA972024 DA972024
469 13 CW067916 104_316_1
474 13 CZ609271 CZ609271 OM_Ba013
476 3 BQ122319 BQ122319 EST607895
476 11 AQ244178 HS_2060_B
477 5 CD743397 IRB3Q3_C0
481 1 AL795151 AL795151
483 10 DR367995 963345 CE
489 11 AQ601600 HS_2105_B
490 11 BH387728 AG-ND-158
490 13 CL323611 RPCI44_45
491 11 AZ431392 IM0216M16
491 14 DE295952 Oryzias_1
493 12 CL721706 OR_BBa005
495 12 CG471620 ZMMBB027
498 11 BH741822 GJC38A04.9
501 11 AQ531090 RPCI-11-3
502 13 CL605608 CH240_182
503 3 BQ305862 MR0-BY200
503 11 AQ426431 C1TBI-E1-
505 1 AI961188 wG59d05.x
505 3 BM586930 BM586930 170006873
505 3 BQ305611 MR0-BY200
510 3 BP068120 BP068120
510 11 AZ031984 RPCI-23-3
513 1 AA390406 LD09377.5
513 5 CD746195 CD746195 S32_G08_S
514 11 AV553180 AV553180 C1TBI-E1-
515 1 BM595704 170006874
516 3 BM641828 170006873
518 3 CD743395 IRB7_D05-
520 5 CD743396 IRB4_D01-
522 5 BG160383 BG160383 df33b05.Y
524 11 AQ520218 HS_5208_A
525 3 BM656275 BM656275 170006873
526 11 BH258725 CH230-75M
527 4 CA938139 say4802.
529 5 CD745002 RB10_C03
530 3 BM629567 170006875
531 5 CD743398 IRB1_A03
531 8 CO193011 EC34984.5
539 3 BM579789 170006872

C 531	15.8	83.2	539	8	CX197198	Sa_mx0_57	604	15.8	83.2	632	13	CW032361	104_261_1
C 532	15.8	83.2	540	3	BM601528	170006870	C 605	15.8	83.2	634	1	AL634018	AL634018
C 533	15.8	83.2	540	3	BM646654	170006873	C 606	15.8	83.2	634	9	DN442792	LI85338-0
C 534	15.8	83.2	541	12	CE416906		C 607	15.8	83.2	635	13	CZ660359	OM_BA021
C 535	15.8	83.2	543	1	AL657096	AL657096	C 608	15.8	83.2	636	8	CN090189	CN090189
C 536	15.8	83.2	543	1	BM650732	170006873	C 609	15.8	83.2	636	8	CN118959	EC2BBA32D
C 537	15.8	83.2	544	7	AW573670	ES7316261	C 610	15.8	83.2	639	1	AL677214	AL677214
C 538	15.8	83.2	546	3	BM598383	170006876	C 611	15.8	83.2	640	1	AL877353	AL877353
C 539	15.8	83.2	547	11	AZ025487		C 612	15.8	83.2	640	4	CA893691	B0180G05-
C 540	15.8	83.2	549	3	BM621991	170006874	C 613	15.8	83.2	641	8	CN086912	EC2BBA27D
C 541	15.8	83.2	551	3	BM623812	170006874	C 614	15.8	83.2	641	8	CO433246	UI-M-HX0-
C 542	15.8	83.2	553	7	AW666072	SK31903.Y	C 615	15.8	83.2	642	3	BQ992129	QGF24M13.
C 543	15.8	83.2	555	7	BF192802	244155.MA	C 616	15.8	83.2	642	13	CL334012	CL334012
C 544	15.8	83.2	557	12	CC951583	BO1D219TF	C 617	15.8	83.2	643	3	CN088569	CN088569
C 545	15.8	83.2	557	13	DU251279	109853818	C 618	15.8	83.2	645	3	BQ937287	AGENCOURT
C 546	15.8	83.2	558	2	BI971816	sa989g11.	C 619	15.8	83.2	645	9	CX994413	Cm_mx0_06
C 547	15.8	83.2	558	4	CB419231	592081.MA	C 620	15.8	83.2	646	9	CX994340	Cm_mx0_05
C 548	15.8	83.2	563	5	CD630861	56067738H	C 621	15.8	83.2	647	1	AL857015	AL857015
C 549	15.8	83.2	563	13	CW760642	OG_BBA006	C 622	15.8	83.2	647	11	AZ995992	2M0282G06
C 550	15.8	83.2	564	8	CV543492	RTS_135.G	C 623	15.8	83.2	647	12	BZ871250	CH240_192
C 551	15.8	83.2	564	11	BH344384		C 624	15.8	83.2	647	14	CR966037	mtH4-JN18
C 552	15.8	83.2	566	3	BM614535		C 625	15.8	83.2	649	1	AL652556	AL652556
C 553	15.8	83.2	567	3	BP743460		C 626	15.8	83.2	649	5	CD630862	CD630862
C 554	15.8	83.2	568	8	CO493668	G.h.fbr-s	C 627	15.8	83.2	650	1	AL648486	AL648486
C 555	15.8	83.2	570	6	CNS095YM		C 628	15.8	83.2	651	12	CE791716	CE791716
C 556	15.8	83.2	570	11	BH382339	AG-ND-162	C 629	15.8	83.2	652	8	CR561186	CR561186
C 557	15.8	83.2	571	8	CO432128	UI-M-HX0-	C 630	15.8	83.2	653	1	AL879054	AL879054
C 558	15.8	83.2	572	13	CL321226		C 631	15.8	83.2	653	3	BQ304357	BQ304357
C 559	15.8	83.2	576	13	CL712193	OR_BBA003	C 632	15.8	83.2	654	5	CF534781	CF534781
C 560	15.8	83.2	577	8	CO983527	GM89019A1	C 633	15.8	83.2	655	2	B1969577	B1969577
C 561	15.8	83.2	577	8	CR441777		C 634	15.8	83.2	655	2	BJ215782	BJ215782
C 562	15.8	83.2	579	8	CA460669	UI-M-HB0-	C 635	15.8	83.2	655	9	CK789708	Sa_mx0_82
C 563	15.8	83.2	580	4	BX269190	BX269190	C 636	15.8	83.2	655	9	DN125520	1139373.M
C 564	15.8	83.2	583	3	BM577967	170006871	C 637	15.8	83.2	655	3	BQ389018	NISC_mq05
C 565	15.8	83.2	583	3	BM646556	170006873	C 638	15.8	83.2	658	3	BQ303735	BQ303735
C 566	15.8	83.2	583	8	CN605889	PEL0408.P	C 639	15.8	83.2	658	5	CF078501	CF078501
C 567	15.8	83.2	583	12	CE248654	tigr-g8s-	C 640	15.8	83.2	658	8	CO045917	UI-M-HS0-
C 568	15.8	83.2	585	4	CB438182		C 641	15.8	83.2	659	8	CV027639	CV027639
C 569	15.8	83.2	585	8	CR417952	CR417952	C 642	15.8	83.2	660	11	BH663938	BH663938
C 570	15.8	83.2	586	3	BM609160	170006870	C 643	15.8	83.2	660	14	DR28865	DR28865
C 571	15.8	83.2	589	3	BM598172	170006876	C 644	15.8	83.2	661	5	AD881756	AD881756
C 572	15.8	83.2	589	3	BU477405	603844352	C 645	15.8	83.2	662	7	AW167183	AW167183
C 573	15.8	83.2	591	3	B0856932	QGB6C10.Y	C 646	15.8	83.2	663	10	DR771540	DR771540
C 574	15.8	83.2	591	6	CNS09FWG		C 647	15.8	83.2	663	13	CZ112701	OM_BA015
C 575	15.8	83.2	591	13	CZ675947		C 648	15.8	83.2	664	12	CE068063	CE068063
C 576	15.8	83.2	593	6	CNS09FWH		C 649	15.8	83.2	665	3	BUI23649	BUI23649
C 577	15.8	83.2	593	14	CR911697	Sus_acrof	C 650	15.8	83.2	665	12	CE023679	CE023679
C 578	15.8	83.2	596	11	AZ222069	RPCI-23-7	C 651	15.8	83.2	665	12	BH980752	BH980752
C 579	15.8	83.2	598	8	CM669755	A0884A06-	C 652	15.8	83.2	666	11	BH980752	BH980752
C 580	15.8	83.2	599	12	CE483243	tigr-g8s-	C 653	15.8	83.2	667	3	BM957610	BM957610
C 581	15.8	83.2	599	14	CR485776		C 654	15.8	83.2	668	8	CN535608	CN535608
C 582	15.8	83.2	600	8	CO429499	UI-M-HW0-	C 655	15.8	83.2	669	5	CK307673	CK307673
C 583	15.8	83.2	600	11	BH561970		C 656	15.8	83.2	670	7	B6123688	B6123688
C 584	15.8	83.2	601	11	BH446690	BOGMJ41TF	C 657	15.8	83.2	670	10	DM126608	DM126608
C 585	15.8	83.2	602	7	AW897906	RC3-NN006	C 658	15.8	83.2	671	12	CE083371	CE083371
C 586	15.8	83.2	605	1	AL959960		C 659	15.8	83.2	672	11	AZ360620	AZ360620
C 587	15.8	83.2	605	5	CK655532		C 660	15.8	83.2	673	12	CE270717	CE270717
C 588	15.8	83.2	607	2	BG712561	pglin.pk0	C 661	15.8	83.2	674	1	AL646912	AL646912
C 589	15.8	83.2	610	13	DU240612		C 662	15.8	83.2	674	8	CR441963	CR441963
C 590	15.8	83.2	614	8	CN086155	EC2BBA26D	C 663	15.8	83.2	675	11	BZ333734	BZ333734
C 591	15.8	83.2	618	10	DY182842	001016BEM	C 664	15.8	83.2	676	7	BZ333734	BZ333734
C 592	15.8	83.2	618	13	CL941999		C 665	15.8	83.2	677	8	CN097884	CN097884
C 593	15.8	83.2	620	13	CL605473		C 666	15.8	83.2	678	10	DM135692	DM135692
C 594	15.8	83.2	622	5	CF355549	V-B-129D1	C 667	15.8	83.2	679	11	BZ433551	BZ433551
C 595	15.8	83.2	622	9	CK736725		C 668	15.8	83.2	681	11	BH966573	BH966573
C 596	15.8	83.2	623	3	BU070367	UI-M-F00-	C 669	15.8	83.2	682	11	BZ088431	BZ088431
C 597	15.8	83.2	623	8	CN076593		C 670	15.8	83.2	683	8	CN078014	CN078014
C 598	15.8	83.2	623	9	CK736706		C 671	15.8	83.2	683	8	CN118958	CN118958
C 599	15.8	83.2	623	11	BH382196		C 672	15.8	83.2	684	11	BZ088475	BZ088475
C 600	15.8	83.2	624	13	CZ636685	OM_BA018	C 673	15.8	83.2	685	1	AL638176	AL638176
C 601	15.8	83.2	629	1	AL877874		C 674	15.8	83.2	685	11	BH276177	BH276177
C 602	15.8	83.2	631	10	DV613107		C 675	15.8	83.2	685	13	CW694856	CW694856
C 603	15.8	83.2	632	11	B70649	CIT-HSP-206	C 676	15.8	83.2	687	4	CA138605	CA138605

677	15.8	83.2	688	14	AG016142	Homo sapi	AG016142	Homo sapi	747	1	AL970811	AL970811
678	15.8	83.2	689	10	DT762365	EST119621	DT762365	EST119621	748	8	CO807632	CO807632
679	15.8	83.2	689	11	BH998479	cei123h05	BH998479	cei123h05	748	11	CK461483	CK461483
680	15.8	83.2	690	11	BH025902	RPCT-24-2	BH025902	RPCT-24-2	748	11	BZ081940	BZ081940
681	15.8	83.2	691	14	AG122398	Pan trogl	AG122398	Pan trogl	748	14	CR294382	CR294382
682	15.8	83.2	693	12	CC743712	ZMMBB011	CC743712	ZMMBB011	749	14	CR455376	CR455376
683	15.8	83.2	695	8	CV177655	LEPISNA45	CV177655	LEPISNA45	749	12	CG559361	CG559361
684	15.8	83.2	696	3	BU208321	603104166	BU208321	603104166	750	4	BW729019	BW729019
685	15.8	83.2	698	9	CK496277	JGI_XZG38	CK496277	JGI_XZG38	750	4	CA765776	CA765776
686	15.8	83.2	698	9	DN267104	L1B30320	DN267104	L1B30320	750	8	CR587686	CR587686
687	15.8	83.2	699	13	CW566004	OA_Aba009	CW566004	OA_Aba009	751	9	DN025001	DN025001
688	15.8	83.2	700	3	BM633530	170006875	BM633530	170006875	751	9	DN025002	DN025002
689	15.8	83.2	700	11	AZ358992	RPCT-23-1	AZ358992	RPCT-23-1	752	8	CR417953	CR417953
690	15.8	83.2	701	12	CC518998	CH240_365	CC518998	CH240_365	753	4	CR417953	CR417953
691	15.8	83.2	702	11	BW759219	BW759219	BW759219	BW759219	753	5	CF783611	CF783611
692	15.8	83.2	702	11	BH936769	cdg52f08	BH936769	cdg52f08	753	11	BZ071125	BZ071125
693	15.8	83.2	702	14	AG178050	Pan trogl	AG178050	Pan trogl	754	14	AG117637	AG117637
694	15.8	83.2	703	12	BZ544887	OGANE607M	BZ544887	OGANE607M	755	8	DN086154	DN086154
695	15.8	83.2	705	8	CN459027	UI-M-HN0-	CN459027	UI-M-HN0-	755	10	DV578484	DV578484
696	15.8	83.2	705	10	DY370977	ZO_Ef000	DY370977	ZO_Ef000	756	10	DV578485	DV578485
697	15.8	83.2	705	11	AZ553100	RPCT-23-2	AZ553100	RPCT-23-2	756	10	DN078013	DN078013
698	15.8	83.2	707	9	DN049795	JGI_CABA5	DN049795	JGI_CABA5	757	8	DN088568	DN088568
699	15.8	83.2	707	14	AG321751	Mus muscu	AG321751	Mus muscu	757	8	CV271341	CV271341
700	15.8	83.2	708	8	CN108252	EC2CA31C	CN108252	EC2CA31C	757	9	CX343624	CX343624
701	15.8	83.2	708	9	DN049796	JGI_CABA5	DN049796	JGI_CABA5	757	9	CX880346	CX880346
702	15.8	83.2	708	12	CL182780	104_394_1	CL182780	104_394_1	757	14	AG525039	AG525039
703	15.8	83.2	709	5	CF783250	AGENCOURT	CF783250	AGENCOURT	758	8	CN536767	CN536767
704	15.8	83.2	710	10	DV188880	BCT032_G0	DV188880	BCT032_G0	758	10	DT428498	DT428498
705	15.8	83.2	711	11	BZ028663	cef01a11	BZ028663	cef01a11	759	4	BX753374	BX753374
706	15.8	83.2	712	14	CT142332	Sus scrofa	CT142332	Sus scrofa	759	4	BX753374	BX753374
707	15.8	83.2	714	5	CD351692	UI-M-G10-	CD351692	UI-M-G10-	759	9	DN084603	DN084603
708	15.8	83.2	715	14	CR346329	metel-8N15	CR346329	metel-8N15	759	10	DT428499	DT428499
709	15.8	83.2	715	13	DU454641	109841579	DU454641	109841579	759	10	DT453795	DT453795
710	15.8	83.2	715	5	CF749061	UI-M-HJ0-	CF749061	UI-M-HJ0-	760	11	BH141740	BH141740
711	15.8	83.2	715	8	CN086911	EC2BBA27D	CN086911	EC2BBA27D	762	5	CJ477741	CJ477741
712	15.8	83.2	715	14	CT268937	Sus scrofa	CT268937	Sus scrofa	762	11	BZ081909	BZ081909
713	15.8	83.2	716	8	CN535770	UI-M-HS0-	CN535770	UI-M-HS0-	765	8	CV271616	CV271616
714	15.8	83.2	717	5	CK654491	AGENCOURT	CK654491	AGENCOURT	765	8	CV271616	CV271616
715	15.8	83.2	717	8	CR586444	CR586444	CR586444	CR586444	766	4	BX731259	BX731259
716	15.8	83.2	717	12	BZ654879	OGANE607C	BZ654879	OGANE607C	769	1	AJ787440	AJ787440
717	15.8	83.2	719	5	CK307034	SB02042A2	CK307034	SB02042A2	769	14	DU830777	DU830777
718	15.8	83.2	719	13	CL814711	OR_CBA003	CL814711	OR_CBA003	770	9	DN020250	DN020250
719	15.8	83.2	720	14	AG515277	Mus muscu	AG515277	Mus muscu	770	12	CC786231	CC786231
720	15.8	83.2	721	18	CN090188	EC2BBA32D	CN090188	EC2BBA32D	773	8	CN308840	CN308840
721	15.8	83.2	721	14	DU575742	OO_Ba006	DU575742	OO_Ba006	773	11	BZ059780	BZ059780
722	15.8	83.2	721	14	DU599467	OO_Ba009	DU599467	OO_Ba009	776	4	BX745148	BX745148
723	15.8	83.2	723	8	CN530632	UI-M-HO0-	CN530632	UI-M-HO0-	776	9	DN084604	DN084604
724	15.8	83.2	723	10	DY353628	ZO_Ec000	DY353628	ZO_Ec000	778	11	BZ120200	BZ120200
725	15.8	83.2	724	4	BX731260	BX731260	BX731260	BX731260	778	12	CC584051	CC584051
726	15.8	83.2	724	5	CK640202	UI-M-HN0-	CK640202	UI-M-HN0-	779	9	CX343625	CX343625
727	15.8	83.2	727	8	CN533965	UI-M-HO0-	CN533965	UI-M-HO0-	782	9	CC756467	CC756467
728	15.8	83.2	729	9	CX998994	JGI_CAAQ2	CX998994	JGI_CAAQ2	782	12	CR580586	CR580586
729	15.8	83.2	729	11	AZ551772	RPCT-23-2	AZ551772	RPCT-23-2	783	8	CU496278	CU496278
730	15.8	83.2	730	13	CL836108	OR_CBA006	CL836108	OR_CBA006	783	13	DU209662	DU209662
731	15.8	83.2	732	11	BH982803	cdff13g09	BH982803	cdff13g09	784	8	CR425988	CR425988
732	15.8	83.2	732	13	CZ359508	ZMMBF0100	CZ359508	ZMMBF0100	784	10	DT453796	DT453796
733	15.8	83.2	735	4	BX696593	BX696593	BX696593	BX696593	784	10	CF151627	CF151627
734	15.8	83.2	735	14	DX162770	OR_Aba001	DX162770	OR_Aba001	785	5	CJ452997	CJ452997
735	15.8	83.2	736	8	CR441778	CR441778	CR441778	CR441778	785	11	BH263821	BH263821
736	15.8	83.2	737	8	CN097883	CR586445	CN097883	CR586445	786	8	CT232101	CT232101
737	15.8	83.2	739	8	DN0599419	OO_Ba009	DN0599419	OO_Ba009	786	14	CT351904	CT351904
738	15.8	83.2	739	14	CU599419	CR580587	CU599419	CR580587	786	12	CC629216	CC629216
739	15.8	83.2	740	8	CR526795	MBM02315	CR526795	MBM02315	790	3	BU121255	BU121255
740	15.8	83.2	740	13	CZ614878	OM_Ba014	CZ614878	OM_Ba014	790	5	CF709562	CF709562
741	15.8	83.2	742	8	CR312500	JGI_XZT85	CR312500	JGI_XZT85	790	8	CR562924	CR562924
742	15.8	83.2	742	14	CR181844	Reverse_6	CR181844	Reverse_6	790	10	DV578476	DV578476
743	15.8	83.2	743	9	CX998993	JGI_CAAQ2	CX998993	JGI_CAAQ2	790	10	BX703018	BX703018
744	15.8	83.2	743	9	DN017894	JGI_CAAQ2	DN017894	JGI_CAAQ2	791	4	BX706436	BX706436
745	15.8	83.2	744	11	BZ031497	ceh43h03	BZ031497	ceh43h03	792	4	CJ474837	CJ474837
746	15.8	83.2	746	8	CN076592	EC2BBA12A	CN076592	EC2BBA12A	792	5	BX703019	BX703019
747	15.8	83.2	746	10	DV574714	0058P0015	DV574714	0058P0015	793	10	DV578482	DV578482
748	15.8	83.2	746	10	DV574715	0058P0015	DV574715	0058P0015	793	10	DV578483	DV578483
749	15.8	83.2	746	10	DV574715	0058P0015	DV574715	0058P0015	793	10	DV578483	DV578483


```
c 969 15..8 83..2 950 7 BE311673
c 970 15..8 83..2 950 14 DX043737
c 971 15..8 83..2 952 11 BH157372
c 972 15..8 83..2 953 14 C9N06HSY
c 973 15..8 83..2 953 14 C9N06HSY
c 974 15..8 83..2 965 4 AG033435
c 975 15..8 83..2 968 4 BY716162
c 976 15..8 83..2 973 13 CL444650
c 977 15..8 83..2 974 3 BU514398
c 978 15..8 83..2 974 5 CF242598
c 979 15..8 83..2 984 1 AV337438
c 980 15..8 83..2 990 6 AY223165
c 981 15..8 83..2 1002 8 CV861525
c 982 15..8 83..2 1004 1 AL568980
c 983 15..8 83..2 1005 1 AM094865
c 984 15..8 83..2 1021 1 AM094863
c 985 15..8 83..2 1028 12 BZ699902
c 986 15..8 83..2 1033 13 CW928899
c 987 15..8 83..2 1036 13 CL235772
c 988 15..8 83..2 1057 3 BU482478
c 989 15..8 83..2 1060 12 CC243593
c 990 15..8 83..2 1064 14 AY404891
c 991 15..8 83..2 1066 13 CL455059
c 992 15..8 83..2 1068 14 C9N05856
c 993 15..8 83..2 1081 13 C2399014
c 994 15..8 83..2 1100 12 CC248668
c 995 15..8 83..2 1101 14 C9N0151E
c 996 15..8 83..2 1104 13 C2965649
c 997 15..8 83..2 1115 14 AG388201
c 998 15..8 83..2 1116 13 CL948053
c 999 15..8 83..2 1119 12 CL021010
c 1000 15..8 83..2 1120 12 CC224627
c 1001 15..8 83..2 1180 11 BZ550135
```

ALIGNMENTS

```
BI063311 158 bp mRNA linear EST 15-JUN-2001
IL3-UT0117-300301-538-E01 UT0117 Homo sapiens cDNA, mRNA sequence.
BI063311
VERSION EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 158)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0117-
300301-538-E01&t3=2001-03-30&t4=1)
Seq primer: puc 18 forward
```

FEATURES

source

High quality sequence stop: 158.

Location/Qualifiers

1..158

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="UT0117"

/note="Organ: uterus_tumor; Vector: puc18; Site:1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) Profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 158;

Best Local Similarity 100.0%; Pred. No. 3e+02; 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0;

QY 1 CTTTACTTCATAGTCCTTTG 19

Db 141 CTTTACTTCATAGTCCTTTG 123

RESULT 2

BE18833

LOCUS 307 bp mRNA linear EST 21-SEP-2000

DEFINITION CM2-BN0302-050700-256-c01 BN0302 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE18833

VERSION BE18833.1 GI:10251067

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominiidae; Homo.

1 (bases 1 to 307)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-BN0302-050

700-256-c01&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 25

High quality sequence stop: 307.

High quality sequence stop: 307.

Location/Qualifiers

1..307

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BN0302"

/note="Organ: breast normal; Vector: puc18; Site:1: SmaI; Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 307;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 201 CTTTACTTCATAGTCTTTG 219

RESULT 3

BE818817
LOCUS CM2-BN0302-050700-256-b07 BN0302 Homo sapiens cDNA, mRNA sequence. EST 21-SEP-2000

DEFINITION BE818817
ACCESSION BE818817
VERSION BE818817.1 GI:10251051
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 341)

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Mateukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=CM2-BN0302-050700-256-b07&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 341.

FEATURES

source

1. 341
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

/note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 341;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 212 CTTTACTTCATAGTCTTTG 230

RESULT 4

LOCUS

BE818831
DEFINITION CM2-BN0302-050700-256-al2 BN0302 Homo sapiens cDNA, mRNA sequence. EST 21-SEP-2000

ACCESSION BE818831

VERSION BE818831.1 GI:10251065

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 365)

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Mateukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=CM2-BN0302-050700-256-al2&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 14

High quality sequence stop: 365.

FEATURES

source

1. 365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

/note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 365;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 226 CTTTACTTCATAGTCTTTG 244

RESULT 5

LOCUS

AA112374
DEFINITION zn68e10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM2-BN0302-050
700-256-clltk3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 384.
High quality sequence stop: 384.

FEATURES

source
1..384
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 216 CTTTACTTCATAGTCTTTG 234
|||||

RESULT 8

BE818828
LOCUS
DEFINITION
CM2-BN0302-050700-256-a05 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE818828
VERSION
BE818828.1 GI:10251062
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE

1 (bases 1 to 396)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM2-BN0302-050

700-256-a05&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 396.
High quality sequence stop: 396.

FEATURES

source
1..396
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 396;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 207 CTTTACTTCATAGTCTTTG 225
|||||

RESULT 9

BE818840
LOCUS
DEFINITION
CM2-BN0302-050700-256-f02 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE818840
VERSION
BE818840.1 GI:10251074
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE

1 (bases 1 to 404)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM2-BN0302-050
700-256-f02&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 404.
High quality sequence stop: 404.

FEATURES

source
1..404
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"

/clone.lib="BN0302"
 /notes="Organ: breast normal; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 404;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 203 CTTTACTTCATAGTCCTTTG 221

RESULT 10

BE818805
 LOCUS 424 bp mRNA linear EST 21-SEP-2000
 DEFINITION CM2-BN0302-040700-253-b10 BN0302 Homo sapiens cDNA, mRNA sequence.
 BE818805
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 1 (bases 1 to 424)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-040

700-253-b10&t3=2000-07-04&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 4
 High quality sequence stop: 424.
 Location/Qualifiers

FEATURES
 source
 1..424
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone.lib="BN0302"
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 449;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 199 CTTTACTTCATAGTCCTTTG 217

RESULT 12

Query Match 100.0%; Score 19; DB 7; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 209 CTTTACTTCATAGTCCTTTG 227

RESULT 11

BE818807
 LOCUS 449 bp mRNA linear EST 21-SEP-2000
 DEFINITION CM2-BN0302-040700-253-d02 BN0302 Homo sapiens cDNA, mRNA sequence.
 BE818807
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 1 (bases 1 to 449)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-040
 700-253-d02&t3=2000-07-04&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 306.
 Location/Qualifiers

FEATURES
 source
 1..449
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone.lib="BN0302"
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 449;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 199 CTTTACTTCATAGTCCTTTG 217

CD672930/c
 LOCUS fg18a05.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
 fgl8a05 5', mRNA sequence.
 ACCESSION CD672930
 VERSION CD672930.1 GI:32174661
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 453)
 AUTHORS Wistow, G., Bernstein, S.L., Ray, S., Wyatt, M.K., Behal, A.,
 Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
 TITLE Expressed sequence tag analysis of adult human iris for the NEIBank
 Project: steroid-response factors and similarities with retinal
 pigment epithelium
 JOURNAL Mol. Vis. 8 (4), 185-195 (2002)
 PUBMED 12107412
 COMMENT Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 18 row: a column: 05
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1. 453
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="fg18a05"
 /tissue_type="iris"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="Human Iris cDNA (Normalized): fg"
 /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
 library (fx) was normalized by self-subtraction. One
 portion of double stranded plasmid DNA representing the
 library was linearized by NotI. This NotI digested library
 was used as a template for biotinylated RNA synthesis
 using SP6 RNA polymerase. Another portion of the double
 stranded plasmid library was converted to single-stranded
 circles in vitro using Gene II and Exonuclease III (Life
 Technologies). Single-stranded DNA (1 mg) was hybridized
 (Cot 500) with 41 mg of Bio-RNA and vector blocking
 oligonucleotides. The hybridized Bio-RNA/ss-circles were
 removed by streptavidin:phenol extraction. EST analysis
 was performed on the library at the NIH Intramural
 Sequencing Center (NISC)."

FEATURES

source

1. 458
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0302"
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from OSESTES PCR (U.S. Letters patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 5; Length 458;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
 |||||

Db 223 CTTTACTTCATAGTCTTTG 241

ORIGIN

Query Match 100.0%; Score 19; DB 5; Length 453;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
 |||||

Db 97 CTTTACTTCATAGTCTTTG 79

RESULT 13

LOCUS

BE818871
 DEFINITION CM2-BN0302-100700-259-f01 BN0302 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE818871
 VERSION BE818871.1 GI:10251105
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 458)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-100
 700-259-f01&t3=2000-07-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 12
 High quality sequence stop: 458.

TITLE

JOURNAL

PUBMED

COMMENT

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM2-BN0302-100
700-259-a05&t3=2000-07-10&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 466.
FEATURES
source
1..466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 100.0%; Score 19; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 226 CTTTACTTCATAGTCCTTTG 244
RESULT 15
BE818852
LOCUS BE818852 467 bp mRNA linear EST 21-SEP-2000
DEFINITION CM2-BN0302-100700-259-a05 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE818852
VERSION BE818852.1 GI:10251086
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 467)
Dias Neto,E., Garcia Correa,R., Vexjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM2-BN0302-100
700-259-a05&t3=2000-07-10&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 467.
FEATURES
source
1..467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 100.0%; Score 19; DB 7; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 205 CTTTACTTCATAGTCCTTTG 223
RESULT 16
BP253378/c
LOCUS BP253378 571 bp mRNA linear EST 15-SEP-2004
DEFINITION BP253378 Sugano cDNA library, Kidney epithelial cell Homo sapiens
cDNA clone HRC03720, mRNA sequence.
ACCESSION BP253378
VERSION BP253378.1 GI:52135659
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 571)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: yusuzuki@hgc.jp
Location/Qualifiers
1..571
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HRC03720"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"
FEATURES
source
1..571
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HRC03720"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"
ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 571;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 563 CTTTACTTCATAGTCTTTG 545

RESULT 17
CB152636/c
LOCUS CB152636 577 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0209868 L12JSHC0s1 Homo sapiens cDNA clone L12JSHC0s1-5-G07
5', mRNA sequence.

ACCESSION CB152636
VERSION CB152636.1 GI:28137590
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 577)

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongseung@mail.kribb.re.kr

Plate: 5, row: G column: 07

High quality sequence stop: 577.

Location/Qualifiers

1. 577

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="L12JSHC0s1-5-G07"

/sex="M"

/cell_line="J-SHC"

/lab_host="Top10P"

/clone_lib="L12JSHC0s1"

/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;

Site 2: NotI; The poly (A) + RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including

EcoRI site by treatment of T4 RNA ligase and the first

strand cDNA was synthesized from oligo dt-selected mRNA by

priming with dt-tailed vector. The dt-tailed vector was

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The

obtained cDNA vectors were used for transformation of

competent cells E. coli Top10P' by electroporation method.

The cDNA libraries constructed by this method are

full-length enriched cDNA library. After analyzing and

sequencing about 2,000 - 3,000 colonies in original cDNA

library, the abundant cDNAs were selected and amplified by

PCR reaction using vector region primer including T7

promotor as 5' primer and N(dt)14 as 3' primer. The PCR

products were used as template for synthesis of

biotinylated single stranded RNA by in vitro transcription

reaction. The synthesized RNA probes were hybridized with

antisense single stranded cDNAs prepared from original

library and incubated with avidin-gel. After removing

DNA-RNA hybrids by centrifuge, the subtracted cDNA

libraries were constructed by transfection of the

ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 577;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 238 CTTTACTTCATAGTCTTTG 220

RESULT 18

BP265234/c

LOCUS BP265234 580 bp mRNA linear EST 16-SEP-2004

DEFINITION BP265234 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone

JTH00322, mRNA sequence.

ACCESSION BP265234

VERSION BP265234.1 GI:52180465

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 580)

Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

Genome Res. 14 (9), 1711-1718 (2004)

15342556

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Tel: 81-3-5449-5343

Fax: 81-3-5449-5416

Email: yuzuki@hgc.jp.

Location/Qualifiers

1. 580

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="JTH00322"

/tissue_type="thyroid"

/cell_line="JTH"

/clone_lib="Sugano cDNA library, thyroid JTH"

/note="thyroid tumor"

ORIGIN

Query Match 100.0%; Score 19; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 368 CTTTACTTCATAGTCTTTG 350

RESULT 19

BP268151/c

LOCUS BP268151 580 bp mRNA linear EST 16-SEP-2004

DEFINITION BP268151 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone

JTH09144, mRNA sequence.

ACCESSION BP268151

VERSION BP268151.1 GI:52183383

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH09144"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/note="thyroid tumor"

ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 579 CTTTACTTCATAGTCTTTG 561
|||||

RESULT 20
BP256144/c
LOCUS BP256144 Sugano cDNA library, kidney epithelial cell Homo sapiens
DEFINITION cDNA clone HRC10671, mRNA sequence.
ACCESSION BP256144
VERSION BP256144.1 GI:52171374
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC10671"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 580;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 579 CTTTACTTCATAGTCTTTG 561
|||||

RESULT 20
BP256147/c
LOCUS BP256147 Sugano cDNA library, kidney epithelial cell Homo sapiens
DEFINITION cDNA clone HRC10676, mRNA sequence.
ACCESSION BP256147
VERSION BP256147.1 GI:52171377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC10671"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 421 CTTTACTTCATAGTCTTTG 403
|||||

Query Match 100.0%; Score 19; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 456 CTTTACTTCATAGTCTTTG 438
|||||

RESULT 21
BP363758/c
LOCUS BP363758 Sugano cDNA library, fetal lung fibroblast TIG Homo sapiens
DEFINITION sapiens cDNA clone TIR02383, mRNA sequence.
ACCESSION BP363758
VERSION BP363758.1 GI:52293963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TIR02383"
/tissue_type="lung"
/cell_type="fibroblast"
/cell_line="TIG"
/dev_stage="fetal"
/clone_lib="Sugano cDNA library, fetal lung fibroblast TIG"

ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 421 CTTTACTTCATAGTCTTTG 403
|||||

RESULT 22
BP256147/c
LOCUS BP256147 Sugano cDNA library, kidney epithelial cell Homo sapiens
DEFINITION cDNA clone HRC10676, mRNA sequence.
ACCESSION BP256147
VERSION BP256147.1 GI:52171377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TIR02383"
/tissue_type="lung"
/cell_type="fibroblast"
/cell_line="TIG"
/dev_stage="fetal"
/clone_lib="Sugano cDNA library, fetal lung fibroblast TIG"

ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
Db 421 CTTTACTTCATAGTCTTTG 403
|||||

RESULT 22
BP256147/c
LOCUS BP256147 Sugano cDNA library, kidney epithelial cell Homo sapiens
DEFINITION cDNA clone HRC10676, mRNA sequence.
ACCESSION BP256147
VERSION BP256147.1 GI:52171377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TIR02383"
/tissue_type="lung"
/cell_type="fibroblast"
/cell_line="TIG"
/dev_stage="fetal"
/clone_lib="Sugano cDNA library, fetal lung fibroblast TIG"

Query Match 100.0%; Score 19; DB 7; Length 630;

Best Local Similarity 100.0%; Pred. No. 3.5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 107 CTTTACTTCATAGTCTTTG 89
|||||

RESULT 25
BM843051/c
LOCUS
DEFINITION K-ES70120649 S12SNU216 Homo sapiens cDNA clone S12SNU216-77-H12 5',
mRNA sequence.

ACCESSION BM843051
VERSION BM843051.1 GI:19199460
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 637)
AUTHORS Oh, J.H., Yang, J.O., Hahn, Y., Kim, M.R., Byun, S.S., Jeon, Y.J.,
Kim, J.M., Song, K.S., Noh, S.M., Kim, S., Yoo, H.S., Kim, Y.S. and
Kim, N.S.

TITLE Transcriptome analysis of human gastric cancer
JOURNAL Mamm. Genome 16 (12), 942-954 (2005)
PUBMED 16341674
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 77 row: H column: 12
High quality sequence stop: 637.
Location/Qualifiers
1. 637

FEATURES
source
1. 637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-77-H12"
/sex="F"
/tissue_type="Lymph node"
/cell_line="SNU-216"
/lab_host="Top10F"
/clone_lib="S12SNU216"
/notes="Organ: Stomach; Vector: pcNS; Site: 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The deacapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 100.0%; Score 19; DB 3; Length 637;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||

Db 100 CTTTACTTCATAGTCTTTG 82

RESULT 26
BF568108/c
LOCUS
DEFINITION

BF568108 699 bp mRNA linear EST 12-DEC-2000
602183908F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4300248 5',
mRNA sequence.

ACCESSION BF568108
VERSION BF568108.1 GI:11641526
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 699)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM158 row: O column: 01
High quality sequence stop: 638.
Location/Qualifiers
1. 699

FEATURES
source
1. 699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300248"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="NIH MGC 42"
/clone_lib="NIH MGC 42"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 7; Length 699;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 362 CTTTACTTCATAGTCTTTG 344
|||||

RESULT 27
BG251131/c
LOCUS

DEFINITION BG251131 702 bp mRNA linear EST 13-FEB-2001
602364932F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473439 5',
mRNA sequence.

ACCESSION BG251131
VERSION BG251131.1 GI:12760947
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 702)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM158 row: O column: 01
High quality sequence stop: 638.
Location/Qualifiers
1. 702

```

REFERENCE
AUTHORS      1 (bases 1 to 702)
TITLE        NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Inyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10295 row: 0 column: 08
High quality sequence stop: 688.
FEATURES
source      Location/Qualifiers
1..702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4473439"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 19; DB 2; Length 702;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 505 CTTTACTTCATAGTCTTTG 487

RESULT 28
BQ887352/c
LOCUS      BQ887352
DEFINITION AGENCOURT 8678064 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381080
5', mRNA sequence.
ACCESSION  BQ887352
VERSION    BQ887352.1 GI:22279366
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE  1 (bases 1 to 851)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2570 row: 1 column: 09
High quality sequence stop: 579.
FEATURES
source      Location/Qualifiers
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

REFERENCE
AUTHORS      1 (bases 1 to 865)
TITLE        NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM14282 row: a column: 21
High quality sequence stop: 886.
FEATURES
source      Location/Qualifiers
1..865
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721245"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN
Query Match      100.0%; Score 19; DB 4; Length 865;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;

```

```

/clone="IMAGE:6381080"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 19; DB 3; Length 851;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 223 CTTTACTTCATAGTCTTTG 205

RESULT 29
CA454746/c
LOCUS      CA454746
DEFINITION AGENCOURT_10763191 MAPcL Homo sapiens cDNA clone IMAGE:6721245 5',
mRNA sequence.
ACCESSION  CA454746
VERSION    CA454746.1 GI:24904781
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE  1 (bases 1 to 865)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM14282 row: a column: 21
High quality sequence stop: 886.
FEATURES
source      Location/Qualifiers
1..865
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721245"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN
Query Match      100.0%; Score 19; DB 4; Length 865;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;

```

```

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 652 CTTTACTTCATAGTCCTTTG 634

RESULT 30
BM011054/c
LOCUS
DEFINITION
  BM011054 877 bp mRNA linear EST 30-OCT-2001
  603634744F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5421755 5',
  mRNA sequence.
ACCESSION
  BM011054
VERSION
  BM011054.1 GI:16525408
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 877)
  NIH-MGC http://mgs.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  JOURNAL
  COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LCM1878 row: h column: 12
  High quality sequence stop: 746.
  Location/Qualifiers
  1..877
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5421755"
  /tissue_type="neuroblastoma, cell line"
  /lab_host="DH10B (phage-resistant)"
  /clone_lib="NIH MGC 47"
  /notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
  EcoRI; cDNA made by oligo-dT priming. Directionally
  cloned into EcoRI/XhoI sites using the following 5',
  adaptor: GGCACGAG(G). Size-selected >500bp for average
  insert size 1.8kb. Library constructed by Ling Hong in
  the laboratory of Gerald M. Rubin (University of
  California, Berkeley) using ZAP-cDNA synthesis kit
  (Stratagene) and Superscript II RT (Life Technologies).
  Note: this is a NIH_MGC Library."

FEATURES
  source

ORIGIN
  Query Match 100.0%; Score 19; DB 2; Length 877;
  Best Local Similarity 100.0%; Pred. No. 3.6e+02;
  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 225 CTTTACTTCATAGTCCTTTG 207

RESULT 31
BQ958341/c
LOCUS
DEFINITION
  BQ958341 899 bp mRNA linear EST 21-AUG-2002
  AGENCOURT_10037048 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6480481
  5', mRNA sequence.
ACCESSION
  BQ958341
VERSION
  BQ958341.1 GI:22373819
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)

```

```

ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 899)
  NIH-MGC http://mgs.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  JOURNAL
  COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-r@mail.nih.gov
  Tissue Procurement: DCTD/DTF
  cDNA Library Preparation: Rubin Laboratory
  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LCM2660 row: f column: 02
  High quality sequence stop: 636.
  Location/Qualifiers
  1..899
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:6480481"
  /tissue_type="carcinoma, cell line"
  /lab_host="DH10B (phage-resistant)"
  /clone_lib="NIH MGC 40"
  /notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
  Site 2: EcoRI; cDNA made by oligo-dT priming.
  Directionally cloned into EcoRI/XhoI sites using the
  following 5' adaptor: GGCACGAG(G). Library constructed by
  Ling Hong in the laboratory of Gerald M. Rubin (University
  of California, Berkeley) using ZAP-cDNA synthesis kit
  (Stratagene) and Superscript II RT (Life Technologies).
  Note: this is a NIH_MGC Library."

ORIGIN
  Query Match 100.0%; Score 19; DB 3; Length 899;
  Best Local Similarity 100.0%; Pred. No. 3.6e+02;
  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 587 CTTTACTTCATAGTCCTTTG 569

RESULT 32
BQ831564/c
LOCUS
DEFINITION
  BQ831564 901 bp mRNA linear EST 22-MAY-2001
  602765819F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4907885 5',
  mRNA sequence.
ACCESSION
  BQ831564
VERSION
  BQ831564.1 GI:14179151
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 901)
  NIH-MGC http://mgs.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  JOURNAL
  COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov

```

Plate: LLCMI813 row: e column: 06
High quality sequence stop: 663.

FEATURES

source Location/Qualifiers

1..901
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4907895"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 901;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 363 CTTTACTTCATAGCTTTG 345

RESULT 33

LOCUS BX372160 918 bp mRNA linear EST 27-APR-2004
DEFINITION BX372160 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK011VF17 3-PRIME, mRNA sequence.

ACCESSION BX372160

VERSION BX372160.1 GI:30452082

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.

1 (bases 1 to 918)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6601.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BA1040ZE07_CS03797_1&c=6601.f

FEATURES

source Location/Qualifiers

1..918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011VF17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 918;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 294 CTTTACTTCATAGCTTTG 312

RESULT 34

LOCUS BU543952/c

DEFINITION BU543952 923 bp mRNA linear EST 13-SEP-2002
AGENCOURT_10333041 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6576152
5', mRNA sequence.

ACCESSION BU543952

VERSION BU543952.1 GI:22854435

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.

1 (bases 1 to 923)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2774 row: d column: 08

High quality sequence stop: 641.

FEATURES

source

1..923

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6576152"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 3; Length 923;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTG 19

Db 587 CTTTACTTCATAGCTTTG 569

RESULT 35

LOCUS BQ685729/c

DEFINITION BQ685729 940 bp mRNA linear EST 15-JUL-2002
AGENCOURT_8344399 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251057
5', mRNA sequence.

```

ACCESSION BQ685729
VERSION BQ685729.1 GI:21811045
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2395 row: b column: 18
High quality sequence stop: 538.
FEATURES
source
Location/Qualifiers
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6251057"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 100.0%; Score 19; DB 3; Length 940;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTG 19
|||||
Db 728 CTTTACTTCATAGTCCTTG 710

RESULT 36
BQ67433/c
LOCUS BQ67433
DEFINITION BQ67433 1023 bp mRNA linear EST 02-APR-2002
AGENCOURT 6758944 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754944
5', mRNA sequence.
ACCESSION BQ67433
VERSION BQ67433.1 GI:19896479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12793 row: c column: 09
High quality sequence stop: 649.
FEATURES
source
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5754944"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SpORF6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is

```

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6601.f

For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS08AG049ZC03_CS04627_16c=6601.f

```

FEATURES
source
Location/Qualifiers
1..962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK011YF17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 100.0%; Score 19; DB 4; Length 962;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTG 19
|||||
Db 289 CTTTACTTCATAGTCCTTG 271

RESULT 37
BQ67433/c
LOCUS BQ67433
DEFINITION BQ67433 1023 bp mRNA linear EST 02-APR-2002
AGENCOURT 6758944 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754944
5', mRNA sequence.
ACCESSION BQ67433
VERSION BQ67433.1 GI:19896479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12793 row: c column: 09
High quality sequence stop: 649.
FEATURES
source
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5754944"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SpORF6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is

```

destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 3; Length 1023;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19

Db 444 CTTTACTTCATAGTCTTTG 426

RESULT 38

CR597125/c

LOCUS

DEFINITION 1830 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CSODK011YF17 of HeLa cells Cot 25-normalized
of Homo sapiens (human).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

Full-length cDNA libraries and normalization

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue

2 (bases 1 to 1830)

Genoscope.

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life technologies, a

division of invitrogen.

Location/Qualifiers

1. 1830

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODK011YF17"

/tissue type="HeLa cells Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 1830;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19

Db 1159 CTTTACTTCATAGTCTTTG 1141

RESULT 39

DU976902

LOCUS

DEFINITION

490 bp DNA linear GSS 03-JAN-2006

AUAC-aaa07d08.g1 Ascaris suum whole genome shotgun library (PMAJ_4

GSS) Ascaris suum genomic, genomic survey sequence.

ACCESSION

VERSION

DU976902

DU976902.1 GI:84308300

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

GSS.

Ascaris suum (pig roundworm)

Ascaris suum

Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida;

Ascaridoidea; Ascarididae; Ascaris.

1 (bases 1 to 490)

Mitreva, M., McCarter, J.P., Ronko, I., Martin, J., Wylie, T., Dante, M.,

Waterston, R.H., Clifton, S.W. and Wilson, R.

Genome Survey sequences from the parasitic nematode Ascaris suum

Unpublished (2005)

Contact: Mitreva M

Washington University in St. Louis

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: nematode@watson.wustl.edu

Genomic DNA provided by Richard E. Davis at the University of

Colorado medical school. DNA sequenced by Washington University

Genome Sequencing Center

Class: shotgun.

FEATURES

source

1. 490

/organism="Ascaris suum"

/mol_type="genomic DNA"

/db_xref="taxon:6253"

/dev_stage="32-64 cell embryos"

/lab_host="GS10"

/clone_lib="Ascaris suum whole genome shotgun library

(PMAJ_4 GSS)"

/note="Vector: pOTW13; Site 1: BstX1; Site 2: BstX1;

Ascaris suum genomic DNA was randomly sheared,

end-repaired and size fractionated to enrich for 6-6.2 kb

fragments. Genomic DNA was provided by Richard E. Davis

at the University of Colorado medical school. Sequencing

by Washington University Genome Sequencing Center, St.

Louis, MO."

ORIGIN

Query Match 94.7%; Score 18; DB 14; Length 490;

Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19

Db 81 TTTACTTCATAGTCTTTG 98

RESULT 40

DU069153/c

LOCUS

DEFINITION

811 bp DNA linear GSS 12-AUG-2005

139562 Tomato HindIII BAC Library Lycopersicon esculentum genomic

clone LE HBa0107M06 5, genomic survey sequence.

ACCESSION

VERSION

DU069153

DU069153.1 GI:72505294

GSS.

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (Solanum lycopersicum)

Eukaryota; Viridiplantae;

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 811)

Mueller, L.A., Buels, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,

Van Eck, J. and Stack, S.

BAC end sequencing from three Solanum lycopersicon libraries

Unpublished (2005)

Other GSSs: 316712

Contact: Lukas Mueller

Tanksley Lab, Dept. of Plant Breeding

Cornell University

251 Emerson Hall, Ithaca, NY 14853, USA

Tel: 607-255-6557

Fax: 607-255-6683

Email: segn-feedback@sgn.cornell.edu
Insert Length: 94875 Std Error: 0.00
Plate: 107 row: M column: 6
Seq primer: T7
Clase: BAC ends
High quality sequence start: 44
High quality sequence stop: 594.
Location/Qualifiers
1. .811
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="LE_HBa0107M06"
/lab_host="E. coli"
/clone_lib="Tomato HindIII BAC Library"
/note="Vector: pBelOBac11; Site_1: HindIII"

ORIGIN

Query Match 94.7%; Score 18; DB 13; Length 811;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTACTTCATAGTCTTTG 19
|||||
Db 765 TTACTTCATAGTCTTTG 748

RESULT 41
CF780868/c
LOCUS CF780868 916 bp mRNA linear EST 20-OCT-2003
DEFINITION AGENCOURT 15939245 NIH_MGC 219 Homo sapiens cDNA clone IMAGE:30523569 5', mRNA sequence.

ACCESSION CF780868
VERSION CF780868.1 GI:37740645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 916)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (fLLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/fLLNL at:
<http://image.llnl.gov>
Plate: NDAM601 row: 1 column: 10
High quality sequence stop: 651.
Location/Qualifiers
1. .916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30523569"
/tissue type="Pooled Chondrosarcoma Tumor cells"
/lab host="DH10B (T1 phage-resistant)"
/clone lib="NIH_MGC 219"
/note="Vector: pYX-Ase; Site_1: EcoRI; Site_2: NotI;
Library is oligo-dT primed and directionally cloned
Denatured RNA was size fractionated on a 1% agarose gel.
First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was

FEATURES source

size selected according tomRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Ase vector. Average insert size
0.5-1kb. Adaptors 5'(AAATTGGCAGGAGG)3' and 5'd
(CCTCGTGCCG)3'. 3' Linker sequence - GGCGCCGTGAGAGCC T18
Sequencing primers 3'end: T3 promoter primer 5'd
(AATTACCTCACTTAAGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Average insert size 2-3kb.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 94.7%; Score 18; DB 5; Length 916;
Best Local Similarity 94.7%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 708 CTTTACTTCATAGTCTTTG 690

RESULT 42
BI193620/c
LOCUS BI193620 1125 bp mRNA linear EST 10-JUL-2001
DEFINITION 602946519F1 NIH_MGC 42 Homo sapiens cDNA clone IMAGE:5089900 5', mRNA sequence.

ACCESSION BI193620
VERSION BI193620.1 GI:14648640
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 1125)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (fLLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/fLLNL at:
<http://image.llnl.gov>
Plate: LLCMI851 row: e column: 05
High quality sequence stop: 529.
Location/Qualifiers
1. .1125
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5089900"
/tissue type="epithelioid carcinoma cell line"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC 42"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCGCCAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

FEATURES source

Query Match 94.7%; Score 18; DB 2; Length 1125;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 2 TTTACTTCATAGTCTTTG 19
Db 334 TTTACTTCATAGTCTTTG 317

RESULT 43
DE194703
LOCUS
DEFINITION Branchiostoma floridae DNA, clone: CH302-076_I05_T7, genomic survey
sequence.
ACCESSION DE194703
VERSION DE194703.1 GI:73603031
KEYWORDS GSS.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Branchiostoma floridae
Eukaryota: Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1
AUTHORS Fujiyama,A., Toyoda,A., Kuroki,Y. and Sakaki,Y.
TITLE BAC end sequences of CHORI-302 Amphioxus Library
JOURNAL Published Only in Database (2005)
REFERENCE 2 (bases 1 to 400)
AUTHORS Fujiyama,A.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-2005) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, 230-0045, Japan
(E-mail:afujiyam@gsc.riken.jp, URL:http://stt.gsc.riken.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
PRIMERS
Sequencing : T7
LIBRARY
Vector : PTARBAC2.1
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. .400
/organism="Branchiostoma floridae"
/mol_type="genomic DNA"
/db_xref="taxon:7739"
/clone="CH302-076_I05_T7"
/sex="male"
/tissue_type="sperm"
/clone_lib="CHORI0302 Amphioxus genomic BAC library"
/notes="Common name:amphioxus"

ORIGIN
Query Match 91.6%; Score 17.4; DB 14; Length 400;
Best Local Similarity 94.7%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 198 CTTTATTTCATAGTCTTTG 216

RESULT 44
AQ435685/c
LOCUS
DEFINITION HS_5148_B1_F10_SP6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=724 Col=19 Row=L, genomic survey sequence.
ACCESSION AQ435685
VERSION AQ435685.1 GI:4547024
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE 1 (bases 1 to 429)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

Qy 2 TTTACTTCATAGTCTTTG 19
Db 334 TTTACTTCATAGTCTTTG 317

RESULT 43
DE194703
LOCUS
DEFINITION Branchiostoma floridae DNA, clone: CH302-076_I05_T7, genomic survey
sequence.
ACCESSION DE194703
VERSION DE194703.1 GI:73603031
KEYWORDS GSS.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Branchiostoma floridae
Eukaryota: Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE 1
AUTHORS Fujiyama,A., Toyoda,A., Kuroki,Y. and Sakaki,Y.
TITLE BAC end sequences of CHORI-302 Amphioxus Library
JOURNAL Published Only in Database (2005)
REFERENCE 2 (bases 1 to 400)
AUTHORS Fujiyama,A.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-2005) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, 230-0045, Japan
(E-mail:afujiyam@gsc.riken.jp, URL:http://stt.gsc.riken.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
PRIMERS
Sequencing : T7
LIBRARY
Vector : PTARBAC2.1
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. .400
/organism="Branchiostoma floridae"
/mol_type="genomic DNA"
/db_xref="taxon:7739"
/clone="CH302-076_I05_T7"
/sex="male"
/tissue_type="sperm"
/clone_lib="CHORI0302 Amphioxus genomic BAC library"
/notes="Common name:amphioxus"

ORIGIN
Query Match 91.6%; Score 17.4; DB 14; Length 400;
Best Local Similarity 94.7%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 198 CTTTATTTCATAGTCTTTG 216

RESULT 44
AQ435685/c
LOCUS
DEFINITION HS_5148_B1_F10_SP6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=724 Col=19 Row=L, genomic survey sequence.
ACCESSION AQ435685
VERSION AQ435685.1 GI:4547024
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE 1 (bases 1 to 429)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

```

```

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 724 row: L column: 19
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 429.
FEATURES
source
1. .429
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=724 Col=19 Row=L"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/notes="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN
Query Match 91.6%; Score 17.4; DB 11; Length 429;
Best Local Similarity 94.7%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 176 CTTTATTTCATAGTCTTTG 158

RESULT 45
AW664308/c
LOCUS
DEFINITION hi09e10.x1 NCI_CGAP_GUI Homo sapiens cDNA clone IMAGE:2971818 3',
mRNA sequence.
ACCESSION AW664308
VERSION AW664308.1 GI:7456849
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE 1 (bases 1 to 498)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@remail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange. The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

```

```

Seq primer: -40UP from Gibco
High quality sequence stop: 406.
Location/Qualifiers
1. .498
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2971818"
/tissue_type="2 pooled high-grade transitional cell
tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GUI"
/notes="Organ: genitourinary tract; Vector: pCMV-SPORT6;
Site 1: Sali; Site 2: Noti; Cloned unidirectionally.
Primer: Oligo dt. - Library constructed by Life
Technologies."

ORIGIN

Query Match 91.6%; Score 17.4; DB 7; Length 498;
Best Local Similarity 94.7%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTTG 19
|||||
Db 211 CTTTACTTCATAGCTCTTTG 193

RESULT 46
LOCUS
DEFINITION v8_p11_p2_triplex_5.1ld_v6 Ambystoma mexicanum cdNA, mRNA
ACCESSION CN050801
VERSION CN050801.1 GI:45821166
KEYWORDS EST.
SOURCE Ambystoma mexicanum (axolotl)
ORGANISM Ambystoma mexicanum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
Ambystoma.
REFERENCE 1 (bases 1 to 525)
Putta, S., Smith, J.J., Walker, J.A., Rondet, M., Weistrock, D.,
Moneghian, J., Samuels, A.K., Kump, K., King, D.C., Maness, N.J.,
Habermann, B., Tanaka, E., Bryant, S.V., Gardiner, D.M., Parichy, D.M.
and Voss, S.R.
From biomedicine to natural history research: EST resources for
ambystomatid salamanders
BMC Genomics 5 (1), 54 (2004)
15310388
Contact: SR Voss
Department of Biology
University of Kentucky
TH Morgan Building, Lexington, KY 40506, USA
Tel: 859 257 9888
Fax: 859 257 1717
Email: srvoos@uky.edu
The EST is quality trimmed at the ends with a 20 base window and
quality threshold of 15 (phred quality score). Please visit
http://salamander.uky.edu For any information (trace, quality files
etc) regarding this EST.
Location/Qualifiers
1. .525
/organism="Ambystoma mexicanum"
/mol_type="mRNA"
/db_xref="taxon:8296"
/tissue_type="Gill Tissue collected from larvae during T4
induced metamorphosis"
/clone_lib="V6"

ORIGIN

Query Match 91.6%; Score 17.4; DB 8; Length 525;
Best Local Similarity 94.7%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

FEATURES
source
1. .525
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0198"
/notes="Organ: placenta normal; Vector: puc18; Site 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 529;
Best Local Similarity 94.7%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTTG 19
|||||
Db 428 CATTACTTCATAGCTCTTTG 446

RESULT 47
LOCUS
DEFINITION RCI-GN0198-011200-013-h01 GN0198 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG009082
VERSION BG009082.1 GI:12454926
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 529)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RCI&t2=RCI-GN0198-
011200-013-h01&t3=2000-12-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 36
High quality sequence stop: 529.

FEATURES
source
1. .529
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0198"
/notes="Organ: placenta normal; Vector: puc18; Site 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 529;
Best Local Similarity 94.7%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTTG 19
|||||
Db 428 CATTACTTCATAGCTCTTTG 446

RESULT 48
LOCUS
DEFINITION DT490210 B21_F11 PT-MB-N-A-15 Populus trichocarpa cDNA clone
DT490210
WS02544.B21_F11 PT-MB-N-A-15 Populus trichocarpa cDNA clone

```

WS02544 F11 3', mRNA sequence.
 DT490210
 VERSION DT490210.1 GI:73887472
 KEYWORDS EST.
 SOURCE ORGANISM Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
 Populus trichocarpa
 Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.
 1 (bases 1 to 565)
 REFERENCE AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
 Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
 Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
 Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
 Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
 Bohlmann,J.
 TITLE The poplar transcriptome: Analysis of expressed sequence tags from
 multiple cDNA libraries
 JOURNAL Unpublished (2004)
 COMMENT Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS02544 row: F column: 11
 High quality sequence stop: 565
 POLYA=Yes.
 FEATURES source Location/Qualifiers
 1..565
 /organism="Populus trichocarpa"
 /mol_type="mRNA"
 /cultivar="Wild clone"
 /db_xref="taxon:3694"
 /clone="WS02544_F11"
 /sex="Male"
 /lab_hosts="E. coli DH10B TI phage resistant cells"
 /clone_lib="PT-MB-N-A-15"
 /notes="Vector: pBluescript II SK (+) XR; Site1: EcoRI (5'
 end of cDNA); Site2: XhoI (3' end of cDNA); Terminal
 vegetative buds from 20 year old trees harvested near
 Corvallis, Oregon on September 19th, 2001. cDNA was
 prepared from 5 micrograms of mRNA and directionally
 ligated into the pBluescript II SK (+) XR vector using the
 pBluescript II XR cDNA Library Construction Kit according
 to manufacturer's instructions with modifications
 (Stratagene). Plasmid DNA was then transformed by
 electroporation into DH10B cells (Invitrogen) for
 propagation. Normalization was applied according to
 published methods [Donaldo M.F. et al. (1996) Genome
 Research 6(9):791] in order to reduce the abundance of
 highly expressed transcripts."
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 10; Length 565;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 139 CTTTACTACATAGTCTTTG 121
 RESULT 49
 BP267627/c
 LOCUS BP267627 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
 DEFINITION JTH07584, mRNA sequence.
 ACCESSION BP267627
 VERSION BP267627.1 GI:52182859
 KEYWORDS EST.
 WS02544 F11 3', mRNA sequence.
 DT490210
 VERSION DT490210.1 GI:73887472
 KEYWORDS EST.
 SOURCE ORGANISM Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
 Populus trichocarpa
 Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.
 1 (bases 1 to 565)
 REFERENCE AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
 Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
 Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
 Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
 Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
 Bohlmann,J.
 TITLE The poplar transcriptome: Analysis of expressed sequence tags from
 multiple cDNA libraries
 JOURNAL Unpublished (2004)
 COMMENT Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS02544 row: F column: 11
 High quality sequence stop: 565
 POLYA=Yes.
 FEATURES source Location/Qualifiers
 1..565
 /organism="Populus trichocarpa"
 /mol_type="mRNA"
 /cultivar="Wild clone"
 /db_xref="taxon:3694"
 /clone="WS02544_F11"
 /sex="Male"
 /lab_hosts="E. coli DH10B TI phage resistant cells"
 /clone_lib="PT-MB-N-A-15"
 /notes="Vector: pBluescript II SK (+) XR; Site1: EcoRI (5'
 end of cDNA); Site2: XhoI (3' end of cDNA); Terminal
 vegetative buds from 20 year old trees harvested near
 Corvallis, Oregon on September 19th, 2001. cDNA was
 prepared from 5 micrograms of mRNA and directionally
 ligated into the pBluescript II SK (+) XR vector using the
 pBluescript II XR cDNA Library Construction Kit according
 to manufacturer's instructions with modifications
 (Stratagene). Plasmid DNA was then transformed by
 electroporation into DH10B cells (Invitrogen) for
 propagation. Normalization was applied according to
 published methods [Donaldo M.F. et al. (1996) Genome
 Research 6(9):791] in order to reduce the abundance of
 highly expressed transcripts."
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 10; Length 565;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 139 CTTTACTACATAGTCTTTG 121
 RESULT 49
 BP267627/c
 LOCUS BP267627 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
 DEFINITION JTH07584, mRNA sequence.
 ACCESSION BP267627
 VERSION BP267627.1 GI:52182859
 KEYWORDS EST.

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 580)
 REFERENCE AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 PUBMED 15342556
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: ysuzuki@hgc.jp.
 FEATURES source Location/Qualifiers
 1..580
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="JTH07584"
 /tissue_type="thyroid"
 /cell_line="JTH"
 /clone_lib="Sugano cDNA library, thyroid JTH"
 /notes="thyroid tumor"
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 3; Length 580;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 580 CTTTACTTCATAGTCTTTG 562
 RESULT 50
 AM058428/c
 LOCUS AM058428 AAFc WHRI BoE01a Brassica oleracea var. alboglabra cDNA
 DEFINITION clone AAFc_WHRI_BoE01a007P16_SP6, mRNA sequence.
 ACCESSION AM058428
 VERSION AM058428.1 GI:75974407
 KEYWORDS EST.
 SOURCE ORGANISM Brassica oleracea var. alboglabra (Chinese kale)
 Brassica oleracea var. alboglabra
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 603)
 REFERENCE AUTHORS Barker,G.C., Sharp,A., Durkin,J., Naylor,R., Edwards,R.,
 Gjetvaj,B., Lydiate,D. and King,G.
 TITLE Analysis of EST libraries of Brassica oleracea prepared from
 differing stages of development
 JOURNAL Unpublished (2005)
 COMMENT Contact: Barker GC
 WHRI
 Warwick University
 Wellesbourne, Warwick, CV35 9EF, UNITED KINGDOM.
 FEATURES source Location/Qualifiers
 1..603
 /organism="Brassica oleracea var. alboglabra"
 /mol_type="mRNA"
 /cultivar="A12"
 /db_xref="taxon:3714"
 /clone="AAFc_WHRI_BoE01a007P16_SP6"
 /tissue_type="etiolated seedling"
 /clone_lib="AAFc_WHRI_BoE01a"
 ORIGIN

Query Match 91.6%; Score 17.4; DB 1; Length 603;
 Best Local Similarity 94.7%; Pred. No. 2e+03; 1; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 ||||| ||||| ||||| |||||
 Db 109 CTTTACTTCATAGTCCTTTG 91

RESULT 51
 AZ353695/c
 LOCUS
 DEFINITION 630 bp DNA linear GSS 02-OCT-2000
 clone UUGC1M0092L09 R, genomic survey sequence.
 AZ353695
 ACCESSION
 VERSION 1 GI:10464470
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 630)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D. Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts
 Unpublished (2000)
 JOURNAL
 COMMENT
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308 Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plates: 0092 row: L column: 09
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 630.

FEATURES
 source
 Location/Qualifiers
 1..630
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0092L09"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: FWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydronamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 [gil4732114|gb|AF129072.1], a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 91.6%; Score 17.4; DB 11; Length 630;
 Best Local Similarity 94.7%; Pred. No. 2e+03; 1; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 ||||| ||||| ||||| |||||
 Db 77 CTTTAAATTCATAGTCCTTTG 59

RESULT 52
 AG158060
 LOCUS
 DEFINITION 635 bp DNA linear GSS 09-JAN-2002
 Pan troglodytes DNA, clone: RP43-023C18.TJ, genomic survey
 sequence.
 AG158060
 ACCESSION
 VERSION 1 GI:16687738
 KEYWORDS
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Pan.

1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library RPCI-43

Unpublished
 2 (bases 1 to 635)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimps@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library RPCI-43. This BAC
 end was generated during the R&D process and may have higher chance
 of clone tracking errors.
 PRIMERS
 Sequencing: TJ

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

Location/Qualifiers

1..635

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="RP43-023C18.TJ"

/sex="male"

/cell_type="lymphocytes"

/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 91.6%; Score 17.4; DB 14; Length 635;
 Best Local Similarity 94.7%; Pred. No. 2e+03; 1; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19

||||| ||||| ||||| |||||

Db 482 CTTTACTTCATAGTCCTTTG 500

RESULT 53

AZ840748

LOCUS

DEFINITION

ACCESSION

AZ840748 644 bp DNA linear GSS 20-FEB-2001
 2M0138K24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC2M0138K24 F, genomic survey sequence.

```

VERSION      AZ840748.1  GI:13010656
KEYWORDS     GSS.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus

REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
              Sciurognathi; Muridae; Murinae; Mus.
              1 (bases 1 to 644)
              Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
              Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
              Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
              Niederhauser, A. and Wright, D., Weiss, R.
              Mouse whole genome scaffolding with paired end reads from 10kb
              plasmid inserts
              Unpublished (2000)
              Contact: Robert B. Weiss
              University of Utah Genome Center
              University of Utah
              Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
              84112, USA
              Tel: 801 585 5606
              Fax: 801 585 7177
              Email: ddu@genetics.utah.edu
              Insert Length: 10000 Std Error: 0.00
              Plate: 0138 row: K column: 24
              Seq primer: CTTGTAAACGACGCCAGT
              Class: plasmid ends
              High quality sequence stop: 644.

FEATURES     source
              1..644
                 /organism="Mus musculus"
                 /mol_type="genomic DNA"
                 /strain="C57BL/6J"
                 /db_xref="taxon:10090"
                 /clones="UUCG2M0138K24"
                 /sex="Male"
                 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                 /clone_lib="Mouse 10kb plasmid UGICM library"
                 /notes="Vector: PWD42nv; Purified genomic DNA from M.
                 musculus C57BL/6J (male) was obtained from the Jackson
                 Laboratory Mouse DNA Resource
                 (http://www.jax.org/resources/documents/dnares/). The DNA
                 was hydrodynamically sheared by repeated passage through a
                 0.005 inch orifice at constant velocity. The sheared DNA
                 was blunt end-repaired with T4 DNA polymerase and T4
                 polynucleotide kinase. Adaptor oligonucleotides were
                 ligated to the blunt ends in high molar excess. The
                 adaptor DNA was purified and size-selected for a 9.5 to
                 10.5 kb range using preparative agarose gel
                 electrophoresis. Vector DNA was prepared from a derivative
                 of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
                 inducible derivative of plasmid R1. The vector was ligated
                 with adaptors complementary to the insert adaptors and
                 purified. The sheared, adaptor mouse DNA was annealed to
                 adaptor vector DNA, and transformed into
                 chemically-competent E. coli XL10-Gold (Stratagene) cells
                 and selected for ampicillin resistance."

ORIGIN
Query Match      91.6%; Score 17.4; DB 11; Length 644;
Best Local Similarity 94.7%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTTG 19
    |||||
Db 156 CTTTCTTCATAGCTCTTTG 174

RESULT 54
DV148009
LOCUS
DEFINITION CV03074B2E03.f1 CV03-normalized library Euphorbia esula cDNA clone
CV03074B2E03.f1 5, mRNA sequence.

```

```

ACCESSION    DV148009
VERSION      DV148009.1  GI:76857016
KEYWORDS     EST.
SOURCE       Euphorbia esula (leafy spurge)
ORGANISM     Euphorbia esula
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
              Euphorbiaceae; Euphorbia.
              1 (bases 1 to 657)
              Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
              Kim, W.R. and Mikel, M.
              Direct submission, Anderson, J.V. 2005
              Unpublished (2005)
              Contact: James V. Anderson
              USDA/ARS
              1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
              58105, USA
              Tel: 701 239 1263
              Fax: 701 239 1252
              Email: andersj@fargo.ars.usda.gov
              Adaptors and tags in 5'-end sequenced clones:
              (Vector) . . . TAAGCTTGATATCG (End Vector) (Start
              EcoRI adaptor) AATTCATTTGTTGGG (End EcoRI adaptor) (Start
              Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TCGGT (End
              Tag) (Start NotI site/Vector) CGCGCCGCCACCGCGG . . . Base
              Calling/Quality Scores: PHRED from Washington University Genome
              Center. Vector Trimming: Cross match from Washington University
              Genome Center PHRAP suite. Low quality bases (phred score < 20)
              were trimmed from both ends of the sequence by an in-house script.
              This sequence is vector free and at least 200 bp in length.

PCR Primers
FORWARD: TAATACGACTCACTATAGG (T7)
BACKWARD: ATTAACCTCTCACTAAG (T3)
Insert Length: 657 Std Error: 0.00
Plate: CV03074B2 row: E column: 03
Seq primer: TAATACGACTCACTATAGG (T7)
High quality sequence stop: 657.

FEATURES     source
              1..657
                 /organism="Euphorbia esula"
                 /mol_type="mRNA"
                 /db_xref="taxon:3993"
                 /clone="CV03074B2E03.f1"
                 /lab_host="DH10B"
                 /clone_lib="CV03-normalized library"
                 /notes="Vector: pBS II SK(+); Site 1: EcoRI (5' side of
                 insert); Site 2: NotI (3' side of insert); Tissue:
                 Seasonal crown buds from outdoor plants, whole plant
                 tissue (leaf, stem, flowers, seeds, shoots, galls, root
                 and crown buds) from outdoor plants, cold acclimated
                 tissues (leaf, stem, and crown buds), crown and root buds
                 from decapitated greenhouse p ants, and seedlings. The
                 library was constructed and normalized as described by
                 Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
                 Research 6(9): 791-806. An identifying tag was added at the
                 3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAATGCGT."

ORIGIN
Query Match      91.6%; Score 17.4; DB 10; Length 657;
Best Local Similarity 94.7%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTTG 19
    |||||
Db 113 CTTTACTTCATATCTTTG 131

RESULT 55
DT482123/c
LOCUS
DEFINITION WS02533_BR_P14 PT-MB-N-A-15 Populus trichocarpa cDNA clone
WS02533_P14 5, mRNA sequence.

```

ACCESSION DT482123
 VERSION DT482123.1 GI:73879385
 SOURCE EST.
 ORGANISM *Populus trichocarpa* (*Populus balsamifera* subsp. *trichocarpa*)
Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; *Populus*.
 1 (bases 1 to 703)
 REFERENCE Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
 Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
 Babakiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A.,
 Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,
 Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and
 Bohlmann, J.
 The poplar transcriptome: Analysis of expressed sequence tags from
 multiple cDNA libraries
 JOURNAL Unpublished (2004)
 COMMENT Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS02533 row: P column: 14
 High quality sequence stop: 703.
 FEATURES source
 1..703
 Location/Qualifiers
 /organism="Populus trichocarpa"
 /mol_type="mRNA"
 /cultivar="Wild clone"
 /db_xref="taxon:3694"
 /clone="WS02533_P14"
 /sex="Male"
 /lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PT-MB-N-A-15"
 /notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
 end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal
 vegetative buds from 20 year old trees harvested near
 Corvallis, Oregon on September 19th, 2001. cDNA was
 prepared from 5 micrograms of mRNA and directionally
 ligated into the pBluescript II SK (+) XR vector using the
 pBluescript II XR cDNA Library Construction Kit according
 to manufacturer's instructions with modifications
 (Stratagene). Plasmid DNA was then transformed by
 electroporation into DH10B cells (Invitrogen) for
 propagation. Normalization was applied according to
 published methods [Bonaldo M.F. et al. (1996) Genome
 Research 6(9):791] in order to reduce the abundance of
 highly expressed transcripts."
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 10; Length 703;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTTTACTTCATAGTCCTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 139 CTTTACTTCATAGTCCTTG 121
 RESULT 56
 BU290329/c 91.6%; Score 17.4; DB 10; Length 703;
 LOCUS DT482123.1 GI:73879385
 DEFINITION 604146187F1 CSEQCHN55 *Gallus gallus* cDNA clone CHST1025n22 5',
 mRNA sequence.
 ACCESSION BU290329
 VERSION BU290329.1 GI:25739785
 KEYWORDS EST.
 SOURCE *Gallus gallus* (chicken)
 ORGANISM *Gallus gallus*
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; *Gallus*.
 1 (bases 1 to 718)
 REFERENCE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 FEATURES source
 1..718
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="ChEST1025n22"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN55"
 /notes="Organ: kidney + adrenal; Vector: pBluescript II
 KS(+); Site 1: EcoRI; Site 2: NotI; This normalized
 library was constructed from 1 million independent clones.
 cDNA synthesis was initiated using an oligo(dT) primer,
 using methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was bluntended, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."
 ORIGIN
 Query Match 91.6%; Score 17.4; DB 3; Length 718;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTTTACTTCATAGTCCTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 350 CTTTACTTCATAGTCCTTG 332
 RESULT 57
 DN485086/c 91.6%; Score 17.4; DB 3; Length 718;
 LOCUS DN485086.3pr *Populus* female catkins cDNA library *Populus trichocarpa*
 DEFINITION M104506.3pr *Populus* female catkins cDNA library *Populus trichocarpa*
 cDNA clone M104E06 3', mRNA sequence.
 ACCESSION DN485086
 VERSION DN485086.1 GI:60694470
 KEYWORDS EST.
 SOURCE *Populus trichocarpa* (*Populus balsamifera* subsp. *trichocarpa*)
Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; *Populus*.
 1 (bases 1 to 735)
 REFERENCE Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,
 Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Fandre, K.,
 Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,
 Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.
 A *Populus* EST resource for plant functional genomics
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (38), 13951-13956 (2004)


```

PUBMED 15353603
COMMENT Other ESTs: M104E06, M104E06.5pr
        Contact: Bo Segerman
        Umea Plant Science Center, Department of Plant Physiology
        Umea University
        901 87 Umea, Sweden
        Tel: +46 90 786 5279
        Fax: +46 90 786 6676
        Email: bo.segerman@plantphys.umu.se.

FEATURES
    source
        1..735
        /organism="Populus trichocarpa"
        /mol_type="mRNA"
        /db_xref="taxon:3694"
        /clone="M104E06"
        /tissue_type="Female catkins"
        /clone_lib="Populus female catkins cDNA library"

ORIGIN
    Query Match      91.6%; Score 17.4; DB 9; Length 735;
    Best Local Similarity 94.7%; Pred. No. 2e+03; 1; Indels 0; Gaps 0;
    Matches 18; Conservative 0; Mismatches 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| |||||
Db 83 CTTTACTACATAGTCTTTG 65

RESULT 58
DV140694
LOCUS DV140694 740 bp mRNA linear EST 03-OCT-2005
DEFINITION CV03126A2C04.f1 CV03-normalized library Euphorbia esula cDNA clone
ACCESSION DV140694
VERSION DV140694.1 GI:76849701
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Euphorbia esula
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
          rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
          Euphorbiaeae; Euphorbia.
REFERENCE 1 (bases 1 to 740)
AUTHORS Anderson,J.V., Horvath,D.P., Thimmapuram,J., Liu,L., Hernandez,A.,
          Kim,W.R. and Mikel,M.
TITLE Direct submission, Anderson,J.V. 2005
JOURNAL Unpublished (2005)
COMMENT Contact: James V. Anderson
          USDA/ARS
          1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
          58105, USA
          Tel: 701 239 1263
          Fax: 701 239 1252
          Email: andersjv@fargo.ars.usda.gov
          Adaptors and tags in 5'-end sequenced clones:
          (Vector). TAGCTTGATATCG (End Vector) (Start
          EcoRI adaptor) AATTCATGTGTTGG (End EcoRI adaptor) (Start
          Tag) (Start NotI site/Vector) GCGCGCCACCGCGG. . . Base
          Calling/Quality Scores: PHRED from Washington University Genome
          Center. Vector Trimming: Cross match from Washington University
          Genome Center PHRAP suite. Low quality bases (Phred score < 20)
          were trimmed from both ends of the sequence by an in-house script.
          This sequence is vector free and at least 200 bp in length.
          PCR Primers
          FORWARD: TAATACGACTCACTATAGG (T7)
          BACKWARD: ATTAACCTCACTAAAG (T3)
          Insert Length: 740 Std Error: 0.00
          Plate: CV03126A2 row: C column: 04
          Seq primer: TAATACGACTCACTATAGG (T7)
          High quality sequence stop: 740.
          Location/Qualifiers
            1..740

```

```

/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03126A2C04.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="vector: pBS II SK(+); Site 1: EcoRI(5' side of
insert); Site 2: NotI (3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGT."

ORIGIN
    Query Match      91.6%; Score 17.4; DB 10; Length 740;
    Best Local Similarity 94.7%; Pred. No. 2e+03; 1; Indels 0; Gaps 0;
    Matches 18; Conservative 0; Mismatches 1;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| |||||
Db 20 CTTTACTTCATATCTTTG 38

RESULT 59
CO081056/c
LOCUS CO081056 765 bp mRNA linear EST 15-JUN-2004
DEFINITION GR_Ea44N24.r GR_Ea Gossypium raimondii cDNA clone GR_Ea44N24 3',
          mRNA sequence.
ACCESSION CO081056
VERSION CO081056.1 GI:48750537
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
          rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 765)
AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
          Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
          Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
          Arizona Genomics Institute
          The University of Arizona
          Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
          Tel: 520 626 9595
          Fax: 520 621 1259
          Email: rwing@genome.arizona.edu
          Plate: 44 row: N column: 24.
          Location/Qualifiers
            1..765
            /organism="Gossypium raimondii"
            /mol_type="mRNA"
            /db_xref="taxon:29730"
            /clone="GR_Ea44N24"
            /tissue_type="whole seedlings"
            /dev_stage="first true leaves"
            /lab_host="DH10B"
            /clone_lib="GR_Ea"
            /note="vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
            EcoRV; Library made by invitrogen with RNA supplied by
            Wendle lab. Directional cloned into NotI-EV. Colonies
            plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
    Query Match      91.6%; Score 17.4; DB 8; Length 765;
    Best Local Similarity 94.7%; Pred. No. 2e+03;

```

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 431 CTTTACTTCATAGTCCTTTG 413

RESULT 60
 CV276848/c
 LOCUS
 DEFINITION
 CV276848 766 bp mRNA linear EST 22-SEP-2004
 cDNA clone WS0141_P08 3', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Populus trichocarpa x Populus deltoides
 Populus trichocarpa x Populus deltoides
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.
 1 (bases 1 to 766)
 REFERENCE
 AUTHORS
 Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
 Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
 Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
 Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
 Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
 Bohlmann,J.
 TITLE
 The poplar transcriptome: Analysis of expressed sequence tags from
 multiple cDNA libraries
 JOURNAL
 COMMENT
 Unpublished (2004)
 Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS0141 row: P column: 08
 High quality sequence stop: 766
 POLYA=Yes.

FEATURES
 source
 1. 766
 /organism="Populus trichocarpa x Populus deltoides"
 /mol_type="mRNA"
 /cultivar="H11-11"
 /db_xref="taxon:3695"
 /clone="WS0141_P08"
 /sex="Male"
 /lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PTXD-IL-A-5"
 /notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Malacosoma disstria Hubner (forest tent caterpillar) mid-instar larvae caged on the sapling using mesh bags. Mature leaves from within the caged region were collected 2 hours, 12 hours, 24 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation."

Query Match 91.6%; Score 17.4; DB 8; Length 766;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN
 Query Match 91.6%; Score 17.4; DB 8; Length 766;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 50 CTTTACTTCATAGTCCTTTG 32

RESULT 61
 DT478689
 LOCUS
 DEFINITION
 DT478689 779 bp mRNA linear EST 29-AUG-2005
 WS02523_BR I07 PT-MB-N-A-15 Populus trichocarpa cDNA clone
 WS02523 I07 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
 Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.
 1 (bases 1 to 779)
 REFERENCE
 AUTHORS
 Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
 Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
 Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
 Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
 Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
 Bohlmann,J.
 TITLE
 The poplar transcriptome: Analysis of expressed sequence tags from
 multiple cDNA libraries
 JOURNAL
 COMMENT
 Unpublished (2004)
 Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS02523 row: I column: 07
 High quality sequence stop: 779
 POLYA=Yes.

FEATURES
 Location/Qualifiers
 1. 779
 /organism="Populus trichocarpa"
 /mol_type="mRNA"
 /cultivar="Wild clone"
 /db_xref="taxon:3694"
 /clone="WS02523_I07"
 /sex="Male"
 /lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PT-MB-N-A-15"
 /notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal vegetative buds from 20 year old trees harvested near Corvallis, Oregon on September 19th, 2001. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Donaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

Query Match 91.6%; Score 17.4; DB 10; Length 779;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN
 Query Match 91.6%; Score 17.4; DB 10; Length 779;
 Best Local Similarity 94.7%; Pred. No. 2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
 |||||
 Db 660 CTTTACTTCATAGTCCTTTG 678

```

RESULT 62
DT483764/c
LOCUS
DEFINITION
    DT483764
    WS02523.B21_I07 PT-MB-N-A-15 Populus trichocarpa cDNA clone
    EST.
ACCESSION
    DT483764
VERSION
    DT483764.1
KEYWORDS
    GI:73881026
SOURCE
    Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
    Populus trichocarpa
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
    rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
    1 (bases 1 to 779)
AUTHORS
    Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
    Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
    Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
    Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
    Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
    Bohlmann,J.
TITLE
    The poplar transcriptome: Analysis of expressed sequence tags from
    multiple cDNA libraries
JOURNAL
    Unpublished (2004)
COMMENT
    Contact: Joerg Bohlmann
    Genome BC forest genomics program
    University of British Columbia
    Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
    Vancouver, British Columbia, Canada, V6T 1Z3
    Tel: 1-604-822-0282
    Fax: 1-604-822-2114
    Email: bohlmann@mel.ubc.ca
    Plate: WS02523 row: I column: 07
    High quality sequence stop: 779
    POLYA=Yes.

FEATURES
    source
    1..779
    Location/Qualifiers
    /organism="Populus trichocarpa"
    /mol_type="mRNA"
    /cultivar="Wild clone"
    /db_xref="taxon:3694"
    /clone="WS02523_I07"
    /sex="Male"
    /lab_host="E. coli DH10B T1 phage resistant cells"
    /clone_lib="PT-MB-N-A-15"
    /notes="vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
    end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal
    vegetative buds from 20 year old trees harvested near
    Corvallis, Oregon on September 19th, 2001. cDNA was
    prepared from 5 micrograms of mRNA and directionally
    ligated into the pBluescript II SK (+) XR vector using the
    pBluescript II XR cDNA Library Construction Kit according
    to manufacturer's instructions with modifications
    (Stratagene). Plasmid DNA was then transformed by
    electroporation into DH10B cells (Invitrogen) for
    propagation. Normalization was applied according to
    published methods [Bonaldo M.F. et al. (1996) Genome
    Research 6(9):791] in order to reduce the abundance of
    highly expressed transcripts."

ORIGIN
    Query Match 91.6%; Score 17.4; DB 10; Length 779;
    Best Local Similarity 94.7%; Pred. No. 2e+03;
    Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTTG 19
    |||||
Db 120 CTTTACTACATAGCTCTTTG 102

RESULT 63
DT483588/c
LOCUS
DEFINITION
    DT483588
    WS02523.B21_A10 PT-MB-N-A-15 Populus trichocarpa cDNA clone
    EST.
ACCESSION
    DT483588
VERSION
    DT483588.1
KEYWORDS
    GI:73880850
SOURCE
    Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
    Populus trichocarpa
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
    rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
    1 (bases 1 to 789)
AUTHORS
    Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
    Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
    Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
    Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
    Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
    Bohlmann,J.
TITLE
    The poplar transcriptome: Analysis of expressed sequence tags from
    multiple cDNA libraries
JOURNAL
    Unpublished (2004)
COMMENT
    Contact: Joerg Bohlmann
    Genome BC forest genomics program
    University of British Columbia
    Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
    Vancouver, British Columbia, Canada, V6T 1Z3
    Tel: 1-604-822-0282
    Fax: 1-604-822-2114
    Email: bohlmann@mel.ubc.ca
    Plate: WS02523 row: A column: 10
    High quality sequence stop: 789
    POLYA=Yes.

FEATURES
    source
    1..789
    Location/Qualifiers
    /organism="Populus trichocarpa"
    /mol_type="mRNA"
    /cultivar="Wild clone"
    /db_xref="taxon:3694"
    /clone="WS02523_A10"
    /sex="Male"
    /lab_host="E. coli DH10B T1 phage resistant cells"
    /clone_lib="PT-MB-N-A-15"
    /notes="vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
    end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal
    vegetative buds from 20 year old trees harvested near
    Corvallis, Oregon on September 19th, 2001. cDNA was
    prepared from 5 micrograms of mRNA and directionally
    ligated into the pBluescript II SK (+) XR vector using the
    pBluescript II XR cDNA Library Construction Kit according
    to manufacturer's instructions with modifications
    (Stratagene). Plasmid DNA was then transformed by
    electroporation into DH10B cells (Invitrogen) for
    propagation. Normalization was applied according to
    published methods [Bonaldo M.F. et al. (1996) Genome
    Research 6(9):791] in order to reduce the abundance of
    highly expressed transcripts."

ORIGIN
    Query Match 91.6%; Score 17.4; DB 10; Length 789;
    Best Local Similarity 94.7%; Pred. No. 2e+03;
    Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTTG 19
    |||||
Db 116 CTTTACTACATAGCTCTTTG 98

RESULT 64
CV258213/c
LOCUS
DEFINITION
    CV258213
    WS0248.B21_O02 PTxD-ICC-N-A-14 Populus trichocarpa x Populus
    deltoides cDNA clone WS0248_O02 3', mRNA sequence.
ACCESSION
    CV258213

```

```

VERSION      CV258213.1  GI:525111188
KEYWORDS     EST.
SOURCE       Populus trichocarpa x Populus deltoides
ORGANISM     Populus trichocarpa x Populus deltoides
              Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              rosids; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE    1 (bases 1 to 802)
AUTHORS      Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
              Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
              Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
              Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
              Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
              Bohlmann,J.
TITLE        The poplar transcriptome: Analysis of expressed sequence tags from
              multiple cDNA libraries
JOURNAL      Unpublished (2004)
COMMENT      Genome BC forest genomics program
              University of British Columbia
              Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
              Vancouver, British Columbia, Canada, V6T 1Z3
              Tel: 1-604-822-0282
              Fax: 1-604-822-2114
              Email: bohlmann@msl.ubc.ca
              Plate: WS0248 row: O column: 02
              High quality sequence stop: 802.
              Location/Qualifiers
FEATURES     source
              1..802
               /organism="Populus trichocarpa x Populus deltoides"
               /mol_type="mRNA"
               /cultivar="Hil-11"
               /db_xrefs="taxon:3695"
               /clone="WS0248_002"
               /sex="Male"
               /lab_host="E. coli DH10B T1 phage resistant cells"
               /clone_lib="PTXD-ICC-N-A-14"
               /notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
               end of cDNA); Site 2: XhoI (3' end of cDNA); Cultured
               cells [de Sa MM et al. (1992) Plant Physiology 98:728-737]
               were grown in media (45mL) supplemented with either 50uM
               salicylic acid, 50uM benzothiadiazole, 50uM methyl
               jasmonate, 20uM chitosan or 200uL of Pollacia radiosa
               extract. Cells were harvested after a 3 hour treatment,
               along with untreated control cells. mRNA was isolated from
               each tissue source independently and equal quantities of
               mRNA from each tissue were then pooled. cDNA was prepared
               from 5 micrograms of mRNA and directionally ligated into
               the pBluescript II SK (+) XR vector using the pBluescript
               II XR cDNA Library Construction Kit according to
               manufacturer's instructions with modifications
               (Stratagene). Plasmid DNA was then transformed by
               electroporation into DH10B cells (Invitrogen) for
               propagation. Normalization was applied according to
               published methods [Ronald M.F. et al. (1996) Genome
               Research 6(9):791] in order to reduce the abundance of
               highly expressed transcripts."

ORIGIN
Query Match      91.6%; Score 17.4; DB 8; Length 802;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db 108 CTTTACTACATGCTTTG 90

RESULT 65
CG329092
LOCUS          830 bp DNA linear GSS 26-AUG-2003
DEFINITION    genomic survey sequence.

ORIGIN
Query Match      91.6%; Score 17.4; DB 8; Length 802;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db 108 CTTTACTACATGCTTTG 90

RESULT 66
CV242157/c
LOCUS          834 bp mRNA linear EST 22-SEP-2004
DEFINITION    WS02514.B21.1 E19 PT-MB-N-A-15 Populus trichocarpa cDNA clone
              WS02514.E19 37, mRNA sequence.

ACCESSION     CV242157
VERSION       CV242157.1 GI:52495132
KEYWORDS      EST.
SOURCE        Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
ORGANISM      Populus trichocarpa
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              rosids; eurosids I; Malpighiales; Salicaceae; Populus.
              1 (bases 1 to 834)
              Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
              Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
              Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
              Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
              Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
              Bohlmann,J.
TITLE        The poplar transcriptome: Analysis of expressed sequence tags from
              multiple cDNA libraries
JOURNAL      Unpublished (2004)
COMMENT      Genome BC forest genomics program
              University of British Columbia
              Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
              Vancouver, British Columbia, Canada, V6T 1Z3
              Tel: 1-604-822-0282

```

Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca
Plate: WS02514 row: E column: 19
High quality sequence stop: 834
POLYA=Yes.

FEATURES

source
Location/Qualifiers
1..834
/organism="Populus trichocarpa"
/mol_type="mRNA"
/cultivar="Wild clone"
/db_xref="taxon:3694"
/clones="WS02514_E19"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PT-MB-N-A-15"
/note="Vector: pBluescript II SK (+) XR; Site1: EcoRI (5' end of cDNA); Site2: XhoI (3' end of cDNA); Terminal vegetative buds from 20 year old trees harvested near Corvallis, Oregon on September 19th, 2001. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 91.6%; Score 17.4; DB 8; Length 834;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
||||| |||||||
Db 120 CTTTACTACATAGTCTTTG 102

RESULT 67
DT506297/c
LOCUS DT506297 837 bp mRNA linear EST 29-AUG-2005
DEFINITION WS0189.C21 D10 PTxD-IL-N-A-9 Populus trichocarpa x Populus deltoides cDNA clone WS0189_D10 3', mRNA sequence.
ACCESSION DT506297
VERSION DT506297.1 GI:73903559
KEYWORDS EST.
SOURCE Populus trichocarpa x Populus deltoides
ORGANISM Populus trichocarpa x Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids 1; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 837)
AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Scott,J., Barber,S., Yang,G., Babakoff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca
Plate: WS0189 row: D column: 10
High quality sequence stop: 837

FEATURES

source
Location/Qualifiers
1..837
/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="Hill-11"
/db_xref="taxon:3695"
/clones="WS0189_D10"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTxD-IL-N-A-9"
/note="Vector: pBluescript II SK (+) XR; Site1: EcoRI (5' end of cDNA); Site2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Malacosoma disstria Hubner (forest tent caterpillar) mid-instar larvae caged on the sapling using mesh bags. Mature leaves from within the caged region were collected 2 hours, 12 hours, 24 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 837;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTACTTCATAGTCTTTG 19
||||| |||||||
Db 315 CTTTACTACATAGTCTTTG 297

RESULT 68
DT487539
LOCUS DT487539 839 bp mRNA linear EST 29-AUG-2005
DEFINITION WS02533.B21_P14 PT-MB-N-A-15 Populus trichocarpa cDNA clone WS02533_P14 3', mRNA sequence.
ACCESSION DT487539
VERSION DT487539.1 GI:73884801
KEYWORDS EST.
SOURCE Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
ORGANISM Populus trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids 1; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 839)
AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Scott,J., Barber,S., Yang,G., Babakoff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca

TITLE The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
JOURNAL Unpublished (2004)
COMMENT Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca

```

Plate: WS02533 row: P column: 14
High quality sequence stop: 839.
Location/Qualifiers
1. .839
/organism="Populus trichocarpa"
/mol_type="mRNA"
/cultivar="Wild clone"
/db_xref="taxon:3694"
/cloneref="WS02533_P14"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clonelib="PT-MB-N-A-15"
/notes="Vector: pBluescript II SK (+) XR; Site_1: EcoRI (5'
end of cDNA); Site_2: XhoI (3' end of cDNA); Terminal
vegetative buds from 20 year old trees harvested near
Corvallis, Oregon on September 19th, 2001. cDNA was
prepared from 5 micrograms of mRNA and directionally
ligated into the pBluescript II SK (+) XR vector using the
pBluescript II XR cDNA Library Construction Kit according
to manufacturer's instructions with modifications
(Stratagene). Plasmid DNA was then transformed by
electroporation into DH10B cells (Invitrogen) for
propagation. Normalization was applied according to
published methods [Bonaldo M.F. et al. (1996) Genome
Research 6(9):791] in order to reduce the abundance of
highly expressed transcripts."

ORIGIN
Query Match 91.6%; Score 17.4; DB 10; Length 839;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTTG 19
||||| |||||||
Db 735 CTTTACTACATGCTCTTG 753

RESULT 69
DT487258/c
LOCUS DT487258 844 bp mRNA linear EST 29-AUG-2005
DEFINITION WS02533 B21_D09 PT-MB-N-A-15 Populus trichocarpa cDNA clone
WS02533_D09 3', mRNA sequence.
ACCESSION DT487258
VERSION DT487258.1 GI:73884520
KEYWORDS EST.
SOURCE Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
ORGANISM Populus trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 844)
Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Maason, A.,
Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,
Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and
Bohlmann, J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca
Plate: WS02533 row: D column: 09
High quality sequence stop: 844
POLYA=Yes. Location/Qualifiers
1. .844

Plate: WS02533 row: P column: 14
High quality sequence stop: 839.
Location/Qualifiers
1. .839
/organism="Populus trichocarpa"
/mol_type="mRNA"
/cultivar="Wild clone"
/db_xref="taxon:3694"
/cloneref="WS02533_P14"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clonelib="PT-MB-N-A-15"
/notes="Vector: pBluescript II SK (+) XR; Site_1: EcoRI (5'
end of cDNA); Site_2: XhoI (3' end of cDNA); Terminal
vegetative buds from 20 year old trees harvested near
Corvallis, Oregon on September 19th, 2001. cDNA was
prepared from 5 micrograms of mRNA and directionally
ligated into the pBluescript II SK (+) XR vector using the
pBluescript II XR cDNA Library Construction Kit according
to manufacturer's instructions with modifications
(Stratagene). Plasmid DNA was then transformed by
electroporation into DH10B cells (Invitrogen) for
propagation. Normalization was applied according to
published methods [Bonaldo M.F. et al. (1996) Genome
Research 6(9):791] in order to reduce the abundance of
highly expressed transcripts."

ORIGIN
Query Match 91.6%; Score 17.4; DB 10; Length 839;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTTG 19
||||| |||||||
Db 735 CTTTACTACATGCTCTTG 753

RESULT 70
CV254849/c
LOCUS CV254849 852 bp mRNA linear EST 22-SEP-2004
DEFINITION WS02411 B21_F21 PTxD-ICC-N-A-14 Populus trichocarpa x Populus
deltoides cDNA clone WS02411_F21 3', mRNA sequence.
ACCESSION CV254849
VERSION CV254849.1 GI:52507824
KEYWORDS EST.
SOURCE Populus trichocarpa x Populus deltoides
ORGANISM Populus trichocarpa x Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 852)
Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Maason, A.,
Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,
Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and
Bohlmann, J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca
Plate: WS02411 row: F column: 21
High quality sequence stop: 852
POLYA=Yes. Location/Qualifiers
1. .852
/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"
/db_xref="taxon:3695"

```

```

/clonew="WS02411_F21"
/sex="Male"
/lab host="E. coli DH10B T1 phage resistant cells"
/clonelib="PTXD-ICC-N-A-14"
/notes="vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Cultured cells (de Sa MM et al. (1992) Plant Physiology 98:728-737) were grown in media (45mL) supplemented with either 50uM salicylic acid, 50uM benzothiadiazole, 50uM methyl jasmonate, 20uM chitosan or 200uL of Pollacia radiosa extract. Cells were harvested after a 3 hour treatment, along with untreated control cells. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

```

ORIGIN

```

Query Match      91.6%; Score 17.4; DB 8; Length 852;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CTTTACTCATAGTCTTTG 19
||||| |||||||
Db 120 CTTTACTCATAGTCTTTG 102

```

RESULT 71

```

DT489385/c
LOCUS      859 bp mRNA linear EST 29-AUG-2005
DEFINITION WS02542 B21 B08 PT-MB-N-A-15 Populus trichocarpa cDNA clone
WS02542 B08 3', mRNA sequence.

```

```

ACCESSION DT489385
VERSION    1
KEYWORDS   GI:73886647
SOURCE     Populus trichocarpa (Populus balsamifera subsp. trichocarpa)

```

ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 859)
Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A.,
Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,
Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and
Bohlmann, J.

```

```

The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)

```

JOURNAL

```

COMMENT    Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca

```

```

Plate: WS02542 row: B column: 08
High quality sequence stop: 859
POLYA=Yes.

```

FEATURES

```

source
1..859
/organism="Populus trichocarpa"
/mol_type="mRNA"

```

```

/cultivar="Wild clone"
/db xref="taxon:3694"
/clonew="WS02542_B08"
/sex="Male"
/lab host="E. coli DH10B T1 phage resistant cells"
/clonelib="PT-MB-N-A-15"
/notes="vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal vegetative buds from 20 year old trees harvested near Corvallis, Oregon on September 19th, 2001. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

```

ORIGIN

```

Query Match      91.6%; Score 17.4; DB 10; Length 859;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CTTTACTCATAGTCTTTG 19
||||| |||||||
Db 121 CTTTACTCATAGTCTTTG 103

```

RESULT 72

```

CV245752/c
LOCUS      871 bp mRNA linear EST 22-SEP-2004
DEFINITION WS0259 B21 E13 PT-MB-N-A-15 Populus trichocarpa cDNA clone
WS0259 E13 3', mRNA sequence.

```

```

ACCESSION CV245752
VERSION    1
KEYWORDS   GI:52498727
SOURCE     Populus trichocarpa (Populus balsamifera subsp. trichocarpa)

```

ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 871)
Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A.,
Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,
Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and
Bohlmann, J.

```

```

The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)

```

JOURNAL

```

COMMENT    Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca

```

```

Plate: WS0259 row: E column: 13
High quality sequence stop: 871
POLYA=Yes.

```

FEATURES

```

source
1..871
/organism="Populus trichocarpa"
/mol_type="mRNA"
/cultivar="Wild clone"
/db xref="taxon:3694"
/clonew="WS0259_E13"
/sex="Male"

```


/lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PT-MB-N-A-15"
 /notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal vegetative buds from 20 year old trees harvested near Corvallis, Oregon on September 19th, 2001. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II SK (+) XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 91.6%; Score 17.4; DB 8; Length 871;
 Best Local Similarity 94.7%; Pred. No. 2.1e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| |||||
 Db 117 CTTTACTACATAGTCTTTG 99

RESULT 73

DT522272/c
 LOCUS DT522272 877 bp mRNA linear EST 30-AUG-2005
 DEFINITION WS02036.B21_I14 PTXN-IB-N-A-11 Populus trichocarpa x Populus nigra cDNA clone WS02036_I14 3', mRNA sequence.

ACCESSION DT522272
 VERSION DT522272.1 GI:73938962

KEYWORDS
 SOURCE Populus trichocarpa x Populus nigra

ORGANISM
 Populus trichocarpa x Populus nigra
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids I; Malpighiales; Salicaceae; Populus.

REFERENCE
 1 (bases 1 to 877)

AUTHORS
 Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.

TITLE
 The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries

JOURNAL
 COMMENT Unpublished (2004)
 Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237, Vancouver, British Columbia, Canada, V6T 1Z3

Tel: 1-604-822-0282
 Fax: 1-604-822-2114

Email: bohlmann@msl.ubc.ca

Plate: WS02036 row: I column: 14

High quality sequence stop: 877

POLYA=Yes.

FEATURES
 source

Location/Qualifiers
 1. .877
 /organism="Populus trichocarpa x Populus nigra"
 /mol_type="mRNA"
 /cultivar="NXM6"
 /db_xref="taxon:293756"
 /clone="WS02036_I14"
 /sex="Male"
 /lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PTXN-IB-N-A-11"
 /notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees

two metres in height and grown under greenhouse conditions were exposed to continuous feeding by *Cryptorhynchus lapathi* (poplar and willow borer) adults caged on the sapling using mesh bags. Bark (with phloem and cambium attached) from within the caged region was harvested 2 hours, 6 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II SK cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 877;
 Best Local Similarity 94.7%; Pred. No. 2.1e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| |||||
 Db 116 CTTTACTACATAGTCTTTG 98

RESULT 74

DT521741/c
 LOCUS DT521741 887 bp mRNA linear EST 30-AUG-2005
 DEFINITION WS02035.B21_B16 PTXN-IB-N-A-11 Populus trichocarpa x Populus nigra cDNA clone WS02035_B16 3', mRNA sequence.

ACCESSION DT521741
 VERSION DT521741.1 GI:73938431

KEYWORDS
 SOURCE Populus trichocarpa x Populus nigra

ORGANISM
 Populus trichocarpa x Populus nigra
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids I; Malpighiales; Salicaceae; Populus.

REFERENCE
 1 (bases 1 to 887)

AUTHORS
 Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.

TITLE
 The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries

JOURNAL

COMMENT Unpublished (2004)

Contact: Joerg Bohlmann

Genome BC forest genomics program

University of British Columbia

Michael Smith Laboratories, 6174 University Boulevard, Rm. 237, Vancouver, British Columbia, Canada, V6T 1Z3

Tel: 1-604-822-0282

Fax: 1-604-822-2114

Email: bohlmann@msl.ubc.ca

Plate: WS02035 row: B column: 16

High quality sequence stop: 887

POLYA=Yes.

FEATURES
 source

Location/Qualifiers
 1. .887
 /organism="Populus trichocarpa x Populus nigra"
 /mol_type="mRNA"
 /cultivar="NXM6"
 /db_xref="taxon:293756"
 /clone="WS02035_B16"
 /sex="Male"
 /lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PTXN-IB-N-A-11"

/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by *Cryptorhynchus lapathi* (poplar and willow borer) adults caged on the sapling using mesh bags. Bark (with phloem and cambium attached) from within the caged region was harvested 2 hours, 6 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 887;
 Best Local Similarity 94.7%; Pred. No. 2.1e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 111 CTTTACTACATAGTCTTTG 93

RESULT 75

CG329079/c
 LOCUS CG329079 897 bp DNA linear GSS 26-AUG-2003
 DEFINITION OG0DS21TH ZM 0.7-1.5 KB Zea mays genomic clone ZMWBMA0695C17,
 genomic survey sequence.

ACCESSION CG329079.1 GI:34246345
 VERSION GSS.
 KEYWORDS Zea mays
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE

AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 TITLE Other_GSSs: OG0DS21TV
 JOURNAL Contact: Cathy Whitelaw
 COMMENT TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@igr.org
 Seq primer: TR
 Class: methylation filtered.

FEATURES

source
 1..897
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMWBMA0695C17"
 /clone_lib="ZM 0.7-1.5 KB"
 /note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 897;
 Best Local Similarity 94.7%; Pred. No. 2.1e+03;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 888 CTTTACTTCAAGTCTTTG 870

RESULT 76

BE548141/c
 LOCUS BE548141 912 bp mRNA linear EST 09-AUG-2000
 DEFINITION 601072396F1 NIH MGC_12 Homo sapiens cDNA clone IMAGE:3458567 5',
 mRNA sequence.

ACCESSION BE548141
 VERSION BE548141.1 GI:9776786
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 912)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-r@mail.nih.gov

REFERENCE

Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LUAM8449 row: p column: 24
 High quality sequence stop: 651.
 Location/Qualifiers
 1..912
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3458567"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

FEATURES

source
 1..912
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3458567"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 91.6%; Score 17.4; DB 7; Length 912;
 Best Local Similarity 94.7%; Pred. No. 2.1e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
 ||||| ||||| ||||| ||||| |||||
 Db 658 CTTGACTTCATAGTCTTTG 640

RESULT 77

CG147825/c
 LOCUS CG147825 955 bp DNA linear GSS 21-AUG-2003
 DEFINITION FUIDS050TD ZM 0.6-1.0 KB Zea mays genomic clone ZMWBTA0557103,
 genomic survey sequence.

ACCESSION CG147825
 VERSION CG147825.1 GI:34038608
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 955)

```

AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Other_GSSs: PUIDG50TB
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TF
              Class: sheared ends.

FEATURES
source
1. .955
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /strain="B73"
   /db_xref="taxon:4577"
   /clone_lib="ZM 0.6 1.0 KB"
   /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
   Cor selected genomic DNA library"

ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 955;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 917 CTTTACTTCAAGTCTTTG 899

RESULT 78
LOCUS      CG330802
DEFINITION CG2BM89TV ZM 0.7 1.5 KB Zea mays genomic clone ZM8BMA0753P09,
            genomic survey sequence.
ACCESSION  CG330802
VERSION     1
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 976)
            Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other_GSSs: CG2BM89TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
            Location/Qualifiers
            1. .976
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone_lib="ZM 0.7 1.5 KB"
               /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
               methylation filtered genomic DNA library"

FEATURES
source
1. .976
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /strain="B73"
   /db_xref="taxon:4577"
   /clone_lib="ZM 0.7 1.5 KB"
   /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
   methylation filtered genomic DNA library"

AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Other_GSSs: PUIDG50TB
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TF
              Class: sheared ends.

FEATURES
source
1. .955
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /strain="B73"
   /db_xref="taxon:4577"
   /clone_lib="ZM 0.6 1.0 KB"
   /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
   Cor selected genomic DNA library"

ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 955;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 917 CTTTACTTCAAGTCTTTG 899

RESULT 78
LOCUS      CG330802
DEFINITION CG2BM89TV ZM 0.7 1.5 KB Zea mays genomic clone ZM8BMA0753P09,
            genomic survey sequence.
ACCESSION  CG330802
VERSION     1
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 976)
            Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other_GSSs: CG2BM89TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
            Location/Qualifiers
            1. .976
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone_lib="ZM 0.7 1.5 KB"
               /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
               methylation filtered genomic DNA library"

FEATURES
source
1. .976
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /strain="B73"
   /db_xref="taxon:4577"
   /clone_lib="ZM 0.7 1.5 KB"
   /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
   methylation filtered genomic DNA library"

AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Other_GSSs: PUIDG50TB
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TF
              Class: sheared ends.

FEATURES
source
1. .955
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /strain="B73"
   /db_xref="taxon:4577"
   /clone_lib="ZM 0.6 1.0 KB"
   /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
   Cor selected genomic DNA library"

ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 976;
Best Local Similarity 94.7%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
    |||||
Db 84 CTTTACTTCAAGTCTTTG 102

RESULT 79
LOCUS      CB828264/c
DEFINITION lJNEST84f6r Lotus japonicus nodule library 5 and 7 week-old Lotus
            japonicus cDNA 5', mRNA sequence.
ACCESSION  CB828264
VERSION     1
KEYWORDS    EST.
SOURCE      Lotus japonicus
ORGANISM    Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
            Lotus.
            1 (bases 1 to 459)
            Colebatch,G., Freund,S., Trevasakis,B and Udvardi,M.
            Lotus japonicus root nodule ESTs: tools for functional genomics
            Unpublished (2000)
            Contact: Udvardi MK
            Molecular Plant Nutrition
            Max Planck Institute of Molecular Plant Physiology
            Am Muehlenberg 1, 14476 Golm, Germany
            Fax: 49 331 567 8250
            Email: udvardi@pimp-golm.mpg.de
            Seq primer: T7
            High quality sequence stop: 459.
            Location/Qualifiers
            1. .459
               /organism="Lotus japonicus"
               /mol_type="mRNA"
               /cultivar="Gifu (B-129)"
               /db_xref="taxon:34305"
               /dev_stage="5 and 7 week-old plants"
               /clone_lib="Lotus japonicus nodule library 5 and 7
               week-old"
               /notes="Organ: Nodule; Vector: pSPORT1; Site 1: SalI;
               Site 2: NotI; The library was prepared using mRNA
               extracted from nodules of 5 and 7 week-old Lotus plants.
               Nodules were induced by, and contained Mesorhizobium
               strain R7A."

FEATURES
source
1. .459
   /organism="Lotus japonicus"
   /mol_type="mRNA"
   /cultivar="Gifu (B-129)"
   /db_xref="taxon:34305"
   /dev_stage="5 and 7 week-old plants"
   /clone_lib="Lotus japonicus nodule library 5 and 7
   week-old"
   /notes="Organ: Nodule; Vector: pSPORT1; Site 1: SalI;
   Site 2: NotI; The library was prepared using mRNA
   extracted from nodules of 5 and 7 week-old Lotus plants.
   Nodules were induced by, and contained Mesorhizobium
   strain R7A."

ORIGIN
Query Match      89.5%; Score 17; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 3e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
    |||||
Db 372 TTACTTCATAGTCTTTG 356

RESULT 80
LOCUS      AQ431031/c
DEFINITION HS 5087_A2 All_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=663 Col=22 Row=A, genomic survey sequence.
ACCESSION  AQ431031
VERSION     1
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 470)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome

JOURNAL
PUBMED
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

COMMENT
10449764
Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center
University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)

or from Research Genetics (info@resgen.com). BAC end Web Server:

http://www.htsc.washington.edu

Plate: 663 row: A column: 22

Seq primer: T7

Class: BAC ends

High quality sequence stop: 470.

Location/Qualifiers

1..470

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="Plate:663 Col=22 Row=A"

/sex="male"

/clone_lib="RPCI-11 Human Male BAC Library"

/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor

and partially digested with a combination of EcoRI and

EcoRI Methylase. Size selected DNA was cloned into the

pBACE3.6 vector at EcoRI sites"

ORIGIN

Query Match 89.5%; Score 17; DB 11; Length 470;

Best Local Similarity 100.0%; Pred. No. 3e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19

|||||

Db 376 TTACTTCATAGTCTTTG 360

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Fax: 49 331 567 8250
Email: udvardi@imp-golm.mpg.de
Seq primer: T7
High quality sequence stop: 552.

FEATURES
source

1..552

/organism="Lotus japonicus"

/mol_type="mRNA"

/cultivar="Gifu (B-129)"

/db_xref="taxon:34305"

/dev_stage="5 and 7 week-old plants"

/clone_lib="Lotus japonicus nodule library 5 and 7

week-old"

/note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI;

Site 2: NotI; The library was prepared using mRNA

extracted from nodules of 5 and 7 week-old Lotus plants.

Nodules were induced by, and contained Mesorhizobium

strain R7A."

ORIGIN

Query Match 89.5%; Score 17; DB 4; Length 552;

Best Local Similarity 100.0%; Pred. No. 3.1e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19

|||||

Db 44 TTACTTCATAGTCTTTG 28

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (G114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 89.5%; Score 17; DB 11; Length 578;
Best Local Similarity 100.0%; Pred. No. 3.1e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGCTTTT 18
|||||
Db 429 TTTACTTCATAGCTTTT 445

RESULT 83

CT259201
LOCUS
DEFINITION
Sus scrofa genomic clone CH242-238M17, genomic survey sequence.
ACCESSION
CT259201
VERSION
CT259201.1 GI:79715797
KEYWORDS
GSS.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 811)

REFERENCE
Humphray, S.J., Plumb, R.W. and Durham, J.L.
AUTHORS
Direct Submission
TITLE
Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished
JOURNAL
This sequence was generated from the SP6 end of BAC 238M17. 238M17 is part of the CHOKI-242 BAC Library created by P. de Jong. Further details: http://www.sanger.ac.uk/Projects/S_scrofa/.

FEATURES

source
1..811
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="CH242-238M17"
/tissue_type="White blood cells"
/notes="vector pTARBAC1.3_BamHI sex female"

ORIGIN

Query Match 89.5%; Score 17; DB 14; Length 811;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGCTTTT 19
|||||
Db 450 TTACTTCATAGCTTTT 466

RESULT 84

BI819265/c

LOCUS
DEFINITION
BI819265
603034634F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175718 5', mRNA sequence.

ACCESSION
BI819265
VERSION
BI819265.1 GI:15930815
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 820)

REFERENCE

AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM11437 row: 1 column: 23
High quality sequence stop: 712.

FEATURES

source
1..820
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5175718"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/notes="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 89.5%; Score 17; DB 2; Length 820;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTT 17
|||||
Db 798 CTTTACTTCATAGCTTT 782

RESULT 85

DU846802
LOCUS
DEFINITION
DU846802
34821 Tomato HindIII BAC Library Lycopersicon esculentum genomic clone LE HBa0106N11 5, genomic survey sequence.

ACCESSION
DU846802.1 GI:84054989

ORIGIN

Query Match 89.5%; Score 17; DB 2; Length 820;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Query Match 89.5%; Score 17; DB 2; Length 820;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 84

BI819265/c

COMMENT

Other GSSs: 34820
 Contact: Lukas Mueller
 Tanksley Lab, Dept. of Plant Breeding
 Cornell University
 251 Emerson Hall, Ithaca, NY 14853, USA
 Tel: 607-255-6557
 Fax: 607-255-6683
 Email: sgn-feedback@sgn.cornell.edu
 Plate: 106 row: N column: 11
 Seq primer: T7
 Class: BAC ends
 High quality sequence start: 29
 High quality sequence stop: 716.

FEATURES

source
 1..884
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="LE_HBa0106N11"
 /lab_host="E. coli"
 /clone_lib="Tomato HindIII BAC Library"
 /note="vector: pBelOBAC11; Site_1: HindIII"

ORIGIN

Query Match 89.5%; Score 17; DB 14; Length 884;
 Best Local Similarity 100.0%; Pred. No. 3.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19

|||||

Db 664 TTACTTCATAGTCCTTTG 680

RESULT 86

DU870629

LOCUS

DEFINITION 99008 Tomato HindIII BAC Library Lycopersicon esculentum genomic clone LE_HBa0132G17 3, genomic survey sequence.

ACCESSION DU870629

VERSION DU870629.1

KEYWORDS GI:84079172

SOURCE GSS.

ORGANISM Lycopersicon esculentum (Solanum lycopersicum)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 894)

REFERENCE Mueller, L.A., Buels, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,

AUTHORS Van Eck, J. and Stack, S.

TITLE BAC end sequencing from three Solanum lycopersicon libraries

JOURNAL Unpublished (2005)

COMMENT Other GSSs: 99009

Contact: Lukas Mueller

Tanksley Lab, Dept. of Plant Breeding

Cornell University

251 Emerson Hall, Ithaca, NY 14853, USA

Tel: 607-255-6557

Fax: 607-255-6683

Email: sgn-feedback@sgn.cornell.edu

Insert Length: 110966 Std Error: 0.00

Plate: 132 row: G column: 17

Seq primer: SP6

Class: BAC ends

High quality sequence start: 21

High quality sequence stop: 396.

FEATURES

source

1..894
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="LE_HBa0132G17"

ORIGIN

Query Match 89.5%; Score 17; DB 14; Length 894;
 Best Local Similarity 100.0%; Pred. No. 3.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19

|||||

Db 24 TTACTTCATAGTCCTTTG 40

RESULT 87

BU931252/c

LOCUS

DEFINITION BU931252 907 bp mRNA linear EST 18-OCT-2002
 AGENCOURT 10478234 NIH_MGC_109 Homo sapiens cDNA clone
 IMAGE:6671391 5', mRNA sequence.

ACCESSION BU931252

VERSION BU931252.1

KEYWORDS GI:24120071

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homniidae; Homo.

1 (bases 1 to 907)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2950 row: d column: 15

High quality sequence stop: 550.

FEATURES

source

1..907
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6671391"
 /tissue_type="teratocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_109"
 /note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 89.5%; Score 17; DB 3; Length 907;
 Best Local Similarity 100.0%; Pred. No. 3.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCCTTT 18

|||||

Db 775 TTACTTCATAGTCCTTT 759

RESULT 88

AK171788/c

LOCUS

DEFINITION AK171788 2921 bp mRNA linear HTC 21-SEP-2005

DEFINITION	Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:F83001E06 product:Transformer-2 protein homolog [Mus musculus], full insert sequence.		
ACCESSION	AK171788		
VERSION	AK171788.1 GI:74152821		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Mus.		
AUTHORS	1 Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
PUBLISHED	10349636		
REFERENCE	2		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBLISHED	11042159		
REFERENCE	3		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,Y., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Onoda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
PUBLISHED	11076861		
REFERENCE	4		
AUTHORS	Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadofa,K., Matsuda,H., Aghburner,M., Batalov,S., Casavant,T., Fleichmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohseuki,S. and Hayashizaki,Y.		
TITLE	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium		
JOURNAL	Functional annotation of a full-length mouse cDNA collection		
PUBLISHED	Nature 409 (6821), 685-690 (2001)		
REFERENCE	5		
AUTHORS	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaide,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragan,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,		

DEFINITION	Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:F83001E06 product:Transformer-2 protein homolog [Mus musculus], full insert sequence.		
ACCESSION	AK171788		
VERSION	AK171788.1 GI:74152821		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Mus.		
AUTHORS	1 Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
PUBLISHED	10349636		
REFERENCE	2		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBLISHED	11042159		
REFERENCE	3		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,Y., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Onoda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
PUBLISHED	11076861		
REFERENCE	4		
AUTHORS	Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadofa,K., Matsuda,H., Aghburner,M., Batalov,S., Casavant,T., Fleichmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohseuki,S. and Hayashizaki,Y.		
TITLE	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium		
JOURNAL	Functional annotation of a full-length mouse cDNA collection		
PUBLISHED	Nature 409 (6821), 685-690 (2001)		
REFERENCE	5		
AUTHORS	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaide,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragan,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,		

Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Frith,M.C.,
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
 Rogers,J., Birney,E. and Hayashizaki,Y.
 FANTOM Consortium

CONSTRM
 TITLE
 Analysis of the mouse transcriptome based on functional annotation

JOURNAL
 PUBMED
 Of 60,770 full-length cDNAs
 Nature 420 (6915), 563-573 (2002)

REFERENCE
 12466851

AUTHORS

6

Maeda,N., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C.,
 Carninci,P., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzius,R.,
 Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R.,
 Zavolan,M., Davis,M.J., Wilming,L.G., Aidinis,V., Allen,J.E.,
 Ambesi-Impombato,A., Apweiler,R., Aturaliya,R.N., Bailey,T.L.,
 Banaal,M., Baxter,L., Beisel,K.W., Bersano,T., Bono,H., Chalk,A.M.,
 Chiu,K.P., Choudhary,V., Christoffels,A., Clutterbuck,D.R.,
 Crowe,M.L., Dalla,E., Dalrymple,B.P., de Bono,B., Della Gatta,G.,
 di Bernardo,D., Down,T., Engstrom,P., Fagiolini,M., Faulkner,G.,
 Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M.,
 Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E.,
 Gustigich,S., Harbers,M., Hayashi,Y., Hensch,T.K., Hirokawa,N.,
 Hill,D., Huminecki,L., Iacono,M., Kawasawa,Y., Kelso,J., Kitamura,H.,
 Jakt,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitamura,H.,
 Kitano,H., Kollas,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K.,
 Kurochkin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J.,
 Liuni,S., McWilliam,S., Madan Babu,M., Madera,M., Marchionni,L.,
 Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S.,
 Morita,K., Mottagui-Tabar,S., Mulder,N., Nakano,N., Nakauchi,H.,
 Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O.,
 Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J., Pavese,G.,
 Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z.,
 Ringwald,M., Rost,B., Ruan,Y., Salzberg,S.L., Sandelin,A.,
 Schneider,C., Schonbach,C., Sekiguchi,K., Semple,C.A., Seno,S.,
 Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D.,
 Sinclair,B., Sperling,S., Stupka,E., Sugita,K., Sultana,R., Taylor,M.S.,
 Takenaka,Y., Taki,K., Tammoja,K., Tan,S.L., Tang,S., Taylor,M.S.,
 Tegner,U., Tschmann,S.A., Ueda,H.R., van Nimwegen,E., Verardo,R.,
 Wei,C.L., Yagi,K., Yamashita,H., Zabarovsky,E., Zhu,S., Zimmer,A.,
 Hide,W., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T.,
 Brusic,V., Quackenbush,J., Wahlestedt,C., Mattick,J.S., Hume,D.A.,
 Kai,C., Sasaki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M.,
 Suzuki,M., Aoki,J., Arakawa,T., Iida,J., Imamura,K., Itoh,M.,
 Kato,T., Kawaji,H., Kawagashira,N., Kawashima,T., Kojima,M.,
 Kondo,S., Konno,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M.,
 Plessy,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K.,
 Watahiki,A., Okamura-Oho,Y., Suzuki,H., Kawai,J. and Hayashizaki,Y.
 FANTOM Consortium

CONSTRM
 TITLE
 The transcriptional landscape of the mammalian genome

JOURNAL
 PUBMED
 Science 309 (5740), 1559-1563 (2005)

REFERENCE
 16141072

AUTHORS

7

Katayama,S., Tomaru,Y., Kasukawa,T., Waki,K., Nakanishi,M.,
 Nakamura,M., Nishida,H., Yap,C.C., Suzuki,M., Kawai,J., Suzuki,H.,
 Carninci,P., Hayashizaki,Y., Wells,C., Frith,M., Ravasi,T.,
 Pang,K.C., Hallinan,J., Mattick,J., Hume,D.A., Lipovich,L.,
 Batalov,S., Engstrom,P.G., Mizuno,Y., Faghihi,M.A., Sandelin,A.,
 Chalk,A.M., Mottagui-Tabar,S., Liang,Z., Lenhard,B. and
 Wahlestedt,C.

CONSTRM
 TITLE
 Antisense transcription in the mammalian transcriptome

JOURNAL
 PUBMED
 Science 309 (5740), 1564-1566 (2005)

REFERENCE
 16141073

AUTHORS

8

(bases 1 to 2923)
 Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K.,
 Hori,F., Iida,J., Imamura,K., Imotani,K., Itoh,M., Kanagawa,S.,
 Kawai,J., Kojima,M., Konno,H., Murata,M., Nakamura,M., Ninomiya,N.,
 Nishiyori,H., Nomura,K., Ohno,M., Sakazume,N., Sano,H., Sasaki,D.,
 Shibata,K., Shiraki,T., Tagami,M., Tagami,Y., Waki,K., Watahiki,A.,
 Muramatsu,M. and Hayashizaki,Y.

Direct Submission

TITLE
 Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa, 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/.

FEATURES

source

1. .2923

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="FANTOM DB:F830206103"

/db_xref="taxon:10090"

/clone="F830206103"

/cissue_type="activated spleen"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

290 .610

/notes="unnamed protein product; Transformer-2 protein

homolog [Mus musculus] (UniProt Q8BVX0, evidence: FASTY,

94.4%ID, 100%length, match=318)

putative"

/codon_start=1

/protein_id="BAE42995.1"

/db_xref="GI:74192357"

/translation="MASGSPCVPRCPSPORSLAADWYLAWSYPLLSLPFLSASHLCS

EVRVGAVGQEAAGSLRPSDGRGFLRWLGPPLLRPSQRLGVRSGRCGCPGGS

LARD"

Query Match 89.5%; Score 17; DB 6; Length 2923;

Best Local Similarity 100.0%; Pred. No. 3.7e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTACTTCATAGTCCTTT 18

|||||

Db 1677 TTACTTCATAGTCCTTT 1661

RESULT 90

AK076606/c

LOCUS

DEFINITION

AK076606

Mus musculus adult male testis cDNA, RIKEN full-length enriched

library, clone:4921532C16 product:unclassified, full insert

sequence.

ACCESSION

AK076606

VERSION

AK076606.1

GI:26096907

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Murinae; Mus.

REFERENCE

1

Carninci,P. and Hayashizaki,Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

PUBMED

10349636

REFERENCE

2

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
PUBMED
REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

TITLE
JOURNAL
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium, the RIKEN Genome Exploration Research Group
Phase I and II Team.

TITLE
JOURNAL
REFERENCE
AUTHORS

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 RIKEN Genome Exploration Research Group, Genome Science Group
(Genome Network Core Team) and the FANTOM Consortium.
Antisense transcription in the Mammalian Transcriptome
Science 309, 1564-1566 (2005)

TITLE
JOURNAL
REFERENCE
AUTHORS

7 The FANTOM Consortium, Riken Genome Exploration Research Group and
Genome Science Group (Genome Network Project Core Group).

TITLE
JOURNAL
REFERENCE
AUTHORS

The Transcriptional Landscape of the Mammalian Genome
Science 309, 1559-1563 (2005)

8 (bases 1 to 2942)
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayaehida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsumura, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp,
URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome-gsc.riken.jp/
URL: http://fantom-gsc.riken.jp/.

Location/Qualifiers

1. .2942
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4921532C16"
/db_xref="taxon:10090"
/clone="4921532C16"

FEATURES
source

/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. .2942
/note="putative
unclassifiable"

misc_feature

ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 2942;
Best Local Similarity 100.0%; Pred. No. 3.7e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 TTTACTTCATAGTCTTT 18

|||||

Db 1707 TTTACTTCATAGTCTTT 1691

RESULT 91

AKI70599/c

LOCUS

AKI70599 3785 bp mRNA linear HTC 21-SEP-2005
Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN
full-length enriched library, clone:F630109013

product:Transformer-2 protein homolog [Mus musculus], full insert
sequence.

ACCESSION

AKI70599 1 GI:74215396

VERSION

AKI70599.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

2 10349636

REFERENCE

3

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

11042159

REFERENCE

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

11076861

REFERENCE

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kado, K., Matsuura, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikolaou, I., Pesole, G.,

Quackenbush, J., Schriml, L. M., Staab, F., Suzuki, R., Tonita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

11217851

5

CONSTRM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Mikado, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

FANTOM Consortium

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420 (6915), 563-573 (2002)

12466851

6

CONSTRM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Belsel, K.W., Bersano, T., Bono, H., Chaik, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminecki, L., Iacono, M., Ikeo, K., Iwano, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Moris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reid, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiyara, K., Saitana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T.,

Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium

The transcriptional landscape of the mammalian genome

Science 309 (5740), 1559-1563 (2005)

16141072

7

CONSTRM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yag, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chaik, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

RIKEN Genome Exploration Research Group

Antisense transcription in the mammalian transcriptome

Science 309 (5740), 1564-1566 (2005)

16141073

8 (bases 1 to 3785)

CONSTRM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. .3785

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="FANTOM_DB:F630109013"

/db_xref="taxon:10090"

/clone="F630109013"

/cell_type="NOD-derived CD11c +ve dendritic cells"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

329. .652

/notes="unlabeled protein product; Transformer-2 protein homolog [Mus musculus] (Uniprot|Q8BVX0, evidence: FASTY, 95.3%ID, 100%length, match=321)

/putative"

/codon_start=1

/protein_id="BAE41904.1"

/db_xref="GI:74215397"

/translation="MASGSPVCPRCTSPQRSLADWYLANWYPLLSFLSASHLC SEVRVPGVQGAAGRSRLRPSDGRCLRWHLGPPPLLRPSGLRVSARGCCPGGE SLARD"

CDS

ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 3785;
 Best Local Similarity 100.0%; Pred. No. 3.8e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
 |||||
 Db 1719 TTTACTTCATAGTCTTT 1703

RESULT 92
 AK076230/c

LOCUS
 DEFINITION AK076230 3796 bp mRNA linear HTC 02-SEP-2005
 Mus musculus 1 day pregnant adult female mammary gland cDNA, RIKEN
 full-length enriched library, clone:4222403J09 product:inferred:
 RIKEN cDNA 1500010G04 gene, full insert sequence.

ACCESSION
 AK076230 GI:26345227

VERSION
 AK076230.1

KEYWORDS
 HTC; CAP trapper.

SOURCE
 Mus musculus (house mouse)

ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
 1 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 10349636

REFERENCE
 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 11042159

REFERENCE
 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitagawa,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsumura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 11076861

REFERENCE
 4 The RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

REFERENCE
 5 The PANTOM Consortium, the RIKEN Genome Exploration Research Group
 Phase I and II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

REFERENCE
 6 RIKEN Genome Exploration Research Group, Genome Science Group
 (Genome Network Core Team) and the PANTOM Consortium.
 Antisense Transcription in the Mammalian Transcriptome
 Science 309, 1564-1566 (2005)

REFERENCE
 7 The PANTOM Consortium, Riken Genome Exploration Research Group and
 Genome Science Group (Genome Network Project Core Group).
 The Transcriptional Landscape of the Mammalian Genome
 Science 309, 1559-1563 (2005)

REFERENCE
 8 (bases 1 to 3796)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
 Hori,P., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,

Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
 Okazaki,Y., Saio,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/.

FEATURES
 source
 1..3796
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:4222403J09"
 /db_xref="taxon:10090"
 /clone="4222403J09"
 /sex="female"
 /tissue_type="mammary gland"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="1 day pregnant adult"
 333..656
 /notes="unnamed protein product; inferred: RIKEN cDNA
 1500010G04 gene (UniGene|Mm.196598, evidence: UG)
 putative"
 /codon_start=1
 /protein_id="BAC36264.1"
 /db_xref="GI:26345228"
 /translatons="MAGSGPCVPCPTSPORSFLAADWYLAWSYPLLSLPFLASHLC
 SEYVRPCAVQGEAAGRSRLRPSDSCRCFLRWLWAPPPLLRPSQRLGKVSARGCCPGGE
 SLVRD"

ORIGIN
 Query Match 89.5%; Score 17; DB 6; Length 3796;
 Best Local Similarity 100.0%; Pred. No. 3.8e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
 |||||
 Db 1727 TTTACTTCATAGTCTTT 1711

RESULT 93
 DE242715/c

LOCUS
 DEFINITION DE242715 91 bp DNA linear GSS 06-JAN-2006
 Trifolium pratense DNA, clone:RCG48925, genomic survey sequence.

ACCESSION
 DE242715

VERSION
 DE242715.1 GI:84463411

KEYWORDS
 GSS.

SOURCE
 Trifolium pratense

ORGANISM
 Trifolium pratense
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Trifolium.
 1
 Sato,S., Isobe,S., Asamizu,E., Nakamura,Y., Ohmido,N., Sakurai,N.,
 Klimentko,I., Sasamoto,S., Wada,T., Watanabe,A. and Tabata,S.

```

TITLE      Comprehensive structural analysis of the genome of red clover
JOURNAL    (Trifolium pratense)
REFERENCE  2 (bases 1 to 91)
AUTHORS    Sato, S.
TITLE      Direct Submission
JOURNAL    Submitted (01-SEP-2005) Shusei Sato, Kazusa DNA Research Institute,
            The First Laboratory for Plant Gene Research; 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 252-0818, Japan
            (E-mail: ssato@kazusa.or.jp, URL: http://www.kazusa.or.jp/,
            Tel: 81-438-52-3935 (ex.2337), Fax: 81-438-52-3934)

FEATURES   Location/Qualifiers
            source
            1..91
                /organism="Trifolium pratense"
                /mol_type="genomic DNA"
                /db_xref="taxon:57577"
                /clone="RCG48925"

ORIGIN
Query Match      86.3%; Score 16.4; DB 14; Length 91;
Best Local Similarity 94.4%; Pred. No. 4.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATGCTCTTT 18
        |||||
Db      29 CTTTACTTCATGCTCTTT 12

RESULT 94
BH497760/c      172 bp      DNA      linear      GSS 13-DEC-2001
LOCUS      BOHPC14TF BOHP Brassica oleracea genomic clone BOHPC14, genomic
DEFINITION      survey sequence.
ACCESSION      BH497760
VERSION      BH497760.1 GI:17705864
KEYWORDS      GSS.
ORGANISM      Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 172)
AUTHORS      Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
            Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
            Whole genome shotgun sequencing of Brassica oleracea and its
            application to gene discovery and annotation in Arabidopsis
            Genome Res. 15 (4), 487-495 (2005)
PUBMED      15805490
COMMENT      Other_GSSs: BOHPC14TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.

FEATURES   Location/Qualifiers
            source
            1..172
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TOL000DH3"
                /db_xref="taxon:3712"
                /clone="BOHPC14"
                /clone_lib="BOHP"
                /notes="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
            genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN
Query Match      86.3%; Score 16.4; DB 11; Length 172;
Best Local Similarity 94.4%; Pred. No. 5.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      2 TTTACTTCATAGCTTTTG 19
        |||||
Db      150 TTTACTTCATAGTTTGTG 133

RESULT 95
BH419641
LOCUS      RIKEN full-length enriched, 12 days embryo spinal cord Mus
DEFINITION      musculus cDNA clone C530017P24 3' similar to Y12883 M.musculus mRNA
            for nonelenium glutathione peroxidase, mRNA sequence.
ACCESSION      BB419641
VERSION      BB419641.1 GI:9240996
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
            1 (bases 1 to 205)
REFERENCE      Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
            Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
            Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
            Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
            Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
            Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
            Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
            Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,
            Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,
            Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A.,
            Watanabe, S., Yamamura, T., Yamataka, I., Yano, R., Yasunishi, A.,
            Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
            Hayashizaki, Y.
            RIKEN Mouse ESTs (Komno, H., et al.)
            Unpublished (2000)
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
            Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
            Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Thermoablation and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
            Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
            Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
            Okazaki, Y. and Hayashizaki, Y.
            Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci, P. and Hayashizaki, Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
            Location/Qualifiers
            1..205
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="C530017P24"
                /tissue type="spinal cord"
                /dev_stage="12 days embryo"
                /lab_host="DH10B"
                /clone_lib="RIKEN full-length enriched, 12 days embryo
            spinal cord"
            /notes="Site 1: Sali; Site 2: BamHI; cDNA library was
            prepared and sequenced in Mouse Genome Encyclopedia
            Project of Genome Exploration Research Group in Riken

```


Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']
 GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5']
 GAGAGAGATTCGAGTTAAATTAATTCCTCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 86.3%; Score 16.4; DB 7; Length 205;
 Best Local Similarity 94.4%; Pred. No. 5.4e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATGAGTCTTT 18
 Db 43 CATTACTTCATGCTTTT 60

RESULT 96

BB171187

LOCUS

DEFINITION BB171187 245 bp mRNA linear EST 29-JUN-2000
 musculus cDNA clone A230025P18 3', mRNA sequence.

ACCESSION BB171187

VERSION BB171187.1

KEYWORDS GI:8830270

SOURCE EST.

ORGANISM Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 245)

Conno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,

Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,

Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,

Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,

Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,

Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,

Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,

Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Yano, R., Yasunishi, A.,

Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A.,

Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and

Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished (2000)

CONTACT Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S.,

Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotabilization and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length

cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,

Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details

FEATURES

source

Location/Qualifiers
 1..245
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="A230025P18"
 /sex="male"
 /tissue_type="hypothalamus"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male hypothalamus"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']
 GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5']
 GAGAGAGATTCGAGTTAAATTAATTCCTCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN
 Query Match 86.3%; Score 16.4; DB 7; Length 245;
 Best Local Similarity 94.4%; Pred. No. 5.4e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

Qy 2 TTTTACTTCATGCTTTTG 19
 Db 122 TTTTCTTCATGCTTTTG 139

RESULT 97

BB176170

LOCUS

DEFINITION BB176170 279 bp mRNA linear EST 29-JUN-2000
 musculus cDNA clone A230060K02 3', mRNA sequence.

ACCESSION BB176170

VERSION BB176170.1

KEYWORDS GI:8835253

SOURCE EST.

ORGANISM Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 279)

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,

Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,

Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,

Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,

Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,

Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,

Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,

Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Yano, R., Yasunishi, A.,

Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A.,

Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and

Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished (2000)

CONTACT Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S.,

Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Tomaru, Y., Carninci, P., Shibata, Y., Iwano, R., Yasunishi, A.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES

Location/Qualifiers
1..279
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A230060K02"
/sex="male"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male
hypothalamus"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN

Query Match 86.3%; Score 16.4; DB 7; Length 279;
Best Local Similarity 94.4%; Pred. No. 5.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TTTACTTCATAGTCTTTG 19
|||||
DB 156 TTTCTTCATAGTCTTTG 173
|||||

RESULT 98

BB027420
LOCUS
DEFINITION BB027420 RIKEN full-length enriched, adult male pituitary gland Mus
musculus cDNA clone 5330437J01 3', mRNA sequence.
BB027420
ACCESSION BB027420.1 GI:8204795
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

REFERENCE

AUTHORS

Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 290)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Sugahara, Y., Suzuki, H., Tagawa, A.,
Takahashi, F., Tomimaga, N., Toy, T., Tsunoda, Y., Watahiki, A.,
Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A.,
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Tomaru, Y., Carninci, P., Shibata, Y., Iwano, R., Yasunishi, A.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES

source

Location/Qualifiers

1..290
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5330437J01"
/sex="male"
/tissue_type="pituitary gland"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male
pituitary gland"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 3.0 and subtraction to Rot = 100.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI."

ORIGIN

Query Match

86.3%; Score 16.4; DB 7; Length 290;

Best Local Similarity 94.4%; Pred. NO. 5.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19
Db 166 TTTTCTTCATAGTCTTTG 183

RESULT 99
CZ443147
LOCUS CZ443147 293 bp mRNA linear GSS 08-APR-2005
DEFINITION CMHD-GT_211D5-3 GTL_R1_pGTIV1 Mus musculus cDNA clone
CMHD-GT_211D5-3 3', mRNA sequence.

ACCESSION CZ443147
VERSION
KEYWORDS GSS.
SOURCE CMHD43147.1 GI:62426451
ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 293)
AUTHORS Sciurognathi; Muroidea; Muridae; Murinae; Mus.
TITLE Stanford, W.L.
JOURNAL www.cmhd.ca
COMMENT Unpublished (2002)
Contact: Stanford WL
Institute of Biomaterials & Biomedical Engineering
University of Toronto
407 Rosebrugh Bldg., 4 Taddle Creek Rd., Toronto, Ontario, Canada
M5S 3G9
Tel: 416 946 8379
Fax: 416 978 4317
Email: william.stanford@utoronto.ca

pgTIV1 Gene trap insertion. The sequence tag is generated by 3' race. The ES cell line harboring this insertion of the target gene is available through the following web site:
http://pokey.ibe.utoronto.ca/sequence_report.php?id=211D5.
Class: Gene Trap.
Location/Qualifiers
1..293
/organism="Mus musculus"
/mol_type="mRNA"
/strain="12983"
/db_xref="taxon:10090"
/clone="CMHD-GT_211D5-3"
/sex="Male"
/cell_type="Embryonic Stem Cell"
/cell_line="R1"
/clone_lib="GTL_R1_pGTIV1"
/note="Vector: pGTIV1"

FEATURES
source

Query Match 86.3%; Score 16.4; DB 13; Length 293;
Best Local Similarity 94.4%; Pred. No. 5.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

Qy 1 CTTTACTTCATAGTCTTT 18
Db 99 CTTTACTTCATAGTCATT 116

RESULT 100
DN761057/c
LOCUS DN761057 300 bp mRNA linear EST 01-JAN-2006
DEFINITION G.hir-stem 3564 PCR amplified cotton stem cDNA Gossypium hirsutum
cDNA, mRNA sequence.

ACCESSION DN761057
VERSION
KEYWORDS DN761057.1 GI:84146815
SOURCE EST.
ORGANISM Gossypium hirsutum (upland cotton)

REFERENCE 1 (bases 1 to 300)
AUTHORS Taliercio, E.
TITLE Cotton ESTs from Stems
JOURNAL Unpublished (2005)
COMMENT Contact: Taliercio, Earl
Earl Taliercio
USDA-ARS
141 Experiment Station Rd, Stoneville, MS 38756, USA
Tel: 6626865363
Fax: 6626865218
Email: ETaliercio@msa-stoneville.ars.usda.gov.
Location/Qualifiers
1..300
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="desl19"
/db_xref="taxon:3635"
/lab_host="E. coli"
/clone_lib="PCR amplified cotton stem cDNA"
/note="Organ: 10 week old stems; Vector: Topo TA; Stems were harvested from field grown plants"

rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 300)
AUTHORS Taliercio, E.
TITLE Cotton ESTs from Stems
JOURNAL Unpublished (2005)
COMMENT Contact: Taliercio, Earl
Earl Taliercio
USDA-ARS
141 Experiment Station Rd, Stoneville, MS 38756, USA
Tel: 6626865363
Fax: 6626865218
Email: ETaliercio@msa-stoneville.ars.usda.gov.
Location/Qualifiers
1..300
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="desl19"
/db_xref="taxon:3635"
/lab_host="E. coli"
/clone_lib="PCR amplified cotton stem cDNA"
/note="Organ: 10 week old stems; Vector: Topo TA; Stems were harvested from field grown plants"

USDA-ARS
141 Experiment Station Rd, Stoneville, MS 38756, USA
Tel: 6626865363
Fax: 6626865218
Email: ETaliercio@msa-stoneville.ars.usda.gov.
Location/Qualifiers
1..300
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="desl19"
/db_xref="taxon:3635"
/lab_host="E. coli"
/clone_lib="PCR amplified cotton stem cDNA"
/note="Organ: 10 week old stems; Vector: Topo TA; Stems were harvested from field grown plants"

FEATURES
source

Query Match 86.3%; Score 16.4; DB 9; Length 300;
Best Local Similarity 94.4%; Pred. No. 5.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

Qy 1 CTTTACTTCATAGTCTTT 18
Db 257 CTTTGCTTCATAGTCTTT 240

Search completed: June 10, 2006, 19:35:23
Job time : 2400.6 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:21:38 ; Search time 42.5917 Seconds

(without alignments)

834.695 Million cell updates/sec

Title: US-10-600-816-32

Perfect score: 19

Sequence: 1 ctttacttcatagtctttg 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Issued Patents NA:*

1: /EMC Celerra_IDS3/ptodata/2/ina/1 COMB.seq.*

2: /EMC Celerra_IDS3/ptodata/2/ina/5 COMB.seq.*

3: /EMC Celerra_IDS3/ptodata/2/ina/6A COMB.seq.*

4: /EMC Celerra_IDS3/ptodata/2/ina/6B COMB.seq.*

5: /EMC Celerra_IDS3/ptodata/2/ina/7 COMB.seq.*

6: /EMC Celerra_IDS3/ptodata/2/ina/H COMB.seq.*

7: /EMC Celerra_IDS3/ptodata/2/ina/PTUS COMB.seq.*

8: /EMC Celerra_IDS3/ptodata/2/ina/PP COMB.seq.*

9: /EMC Celerra_IDS3/ptodata/2/ina/RE COMB.seq.*

10: /EMC Celerra_IDS3/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	100.0	601	3	US-09-949-016-29173
C 2	19	100.0	601	3	US-09-949-016-126957
C 3	19	100.0	1212	3	US-09-188-930-249
C 4	19	100.0	1212	3	US-09-312-283C-249
C 5	19	100.0	1228	3	US-09-495-050A-223
C 6	19	100.0	2289	3	US-09-949-016-3613
C 7	19	100.0	2456	3	US-09-949-016-572
C 8	19	100.0	26086	3	US-09-949-016-15355
C 9	19	100.0	26238	3	US-09-949-016-12314
C 10	17.4	91.6	601	3	US-09-949-016-173650
C 11	17.4	91.6	50950	3	US-09-949-016-16659
C 12	16.4	86.3	407	3	US-09-513-999C-32761
C 13	16.4	86.3	15131	3	US-09-949-016-16805
C 14	16.4	86.3	22614	3	US-09-949-016-12148
C 15	16.4	86.3	22615	3	US-09-949-016-16745
C 16	16.4	86.3	107937	3	US-09-949-016-17192
C 17	16.4	86.3	139049	3	US-09-949-016-17030
C 18	16	84.2	2363	3	US-09-949-016-2159
C 19	16	84.2	22266	3	US-09-949-016-13901
C 20	16	84.2	130563	3	US-09-949-016-12773
C 21	16	84.2	131379	3	US-09-949-016-16050
C 22	15.8	83.2	601	3	US-09-949-016-48463
C 23	15.8	83.2	601	3	US-09-949-016-48464

Sequence 48465, A
Sequence 48466, A
Sequence 48467, A
Sequence 48468, A
Sequence 121554,
Sequence 12078, A
Sequence 1023, Ap
Sequence 15279, A
Sequence 15133, A
Sequence 15990, A
Sequence 13175, A
Sequence 12009, A
Sequence 17053, A
Sequence 16435, A
Sequence 154, App
Sequence 141, App
Sequence 21055, A
Sequence 21056, A
Sequence 21057, A
Sequence 21058, A
Sequence 60155, A
Sequence 77977, A
Sequence 77978, A
Sequence 77979, A
Sequence 77980, A
Sequence 144809,
Sequence 162942,
Sequence 8840, Ap
Sequence 382, App
Sequence 659, App
Sequence 295, App
Sequence 295, App
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 30, Appli
Sequence 41, Appli
Sequence 38, Appli
Sequence 1289, Ap
Sequence 2698, Ap
Sequence 6, Appli
Sequence 16, Appli
Sequence 16, Appli
Patent No. 5455158
Sequence 1, Appli
Sequence 1, Appli
Sequence 135, App
Sequence 2899, Ap
Sequence 16147, A
Sequence 11904, A
Sequence 14016, A
Sequence 15836, A
Sequence 13674, A
Sequence 13242, A
Sequence 574, App
Sequence 802, App
Sequence 16480, A
Sequence 14868, A
Sequence 14588, A
Sequence 1589, A
Sequence 17119, A
Sequence 13498, A
Sequence 14543, A
Sequence 16297, A
Sequence 13557, A
Sequence 211, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
US-09-949-016-48465
US-09-949-016-48466
US-09-949-016-48467
US-09-949-016-48468
US-09-949-016-121554
US-09-270-767-12078
US-09-620-312D-1023
US-09-270-767-15279
US-09-949-016-15133
US-09-949-016-15990
US-09-949-016-13175
US-09-949-016-12009
US-09-949-016-17053
US-09-949-016-16435
US-09-248-796A-154
US-09-397-787-141
US-09-949-016-21055
US-09-949-016-21056
US-09-949-016-21057
US-09-949-016-21058
US-09-949-016-60155
US-09-949-016-77977
US-09-949-016-77978
US-09-949-016-77979
US-09-949-016-77980
US-09-949-016-144809
US-09-949-016-162942
US-09-949-002-8840
US-09-614-221A-382
US-10-104-047-659
US-08-956-171B-295
US-08-781-986A-295
US-09-248-528-5
US-09-549-108-5
US-09-549-111-5
US-09-550-106-5
US-09-550-394-5
US-09-354-147C-6
US-08-927-539A-30
US-09-354-147C-41
US-09-220-132-38
US-09-023-655-1289
US-09-543-679A-2698
PCT-US95-09819-6
US-08-259-569-16
US-08-826-885-16
5455158-2
US-08-551-356-1
PCT-US93-12687-1
US-09-566-921-135
US-09-543-679A-2699
US-09-949-016-16147
US-09-949-016-11904
US-09-949-016-14016
US-09-949-016-15836
US-09-949-016-12674
US-09-949-016-13242
US-09-949-002-574
US-09-949-002-802
US-09-949-016-16480
US-09-949-016-11868
US-09-949-016-14588
US-09-949-016-14589
US-09-949-016-17119
US-09-949-016-13498
US-09-949-016-14543
US-09-949-016-16297
US-09-949-016-12557
US-09-531-120-211
US-08-916-421B-1
US-09-692-570-1
US-08-989-299-3
US-09-407-427-3

C 97	15	78.9	2415	3	US-09-635-501-3	Sequence 3, Appli	C 170	14.8	77.9	3111	3	US-09-504-393-2	Sequence 2, Appli
C 98	15	78.9	2415	4	US-09-163-648-3	Sequence 3, Appli	C 171	14.8	77.9	3111	3	US-10-053-192-2	Sequence 2, Appli
C 99	15	78.9	2920	3	US-10-158-847-137	Sequence 137, App	C 172	14.8	77.9	3177	3	US-09-248-796A-5784	Sequence 5294, App
C 100	15	78.9	2920	3	US-10-158-825-137	Sequence 137, App	C 173	14.8	77.9	4146	3	US-09-614-221A-229	Sequence 2729, App
C 101	15	78.9	3388	3	US-09-949-016-5413	Sequence 5413, Ap	C 174	14.8	77.9	7680	3	US-09-210-748A-3	Sequence 3, Appli
C 102	15	78.9	3386	3	US-08-989-299-1	Sequence 1, Appli	C 175	14.8	77.9	7680	3	US-09-939-581A-3	Sequence 3, Appli
C 103	15	78.9	3386	3	US-10-158-847-141	Sequence 141, App	C 176	14.8	77.9	10034	3	US-09-771-357-102	Sequence 102, App
C 104	15	78.9	3386	3	US-09-407-427-1	Sequence 1, Appli	C 177	14.8	77.9	10034	3	US-10-059-579A-102	Sequence 102, App
C 105	15	78.9	3396	3	US-09-635-501-1	Sequence 1, Appli	C 178	14.8	77.9	11411	3	US-09-949-016-13357	Sequence 13357, A
C 106	15	78.9	3396	3	US-10-158-825-141	Sequence 141, App	C 179	14.8	77.9	15402	3	US-09-949-016-12868	Sequence 12868, A
C 107	15	78.9	3396	4	US-09-163-648-1	Sequence 1, Appli	C 180	14.8	77.9	17842	3	US-09-949-016-13826	Sequence 13826, A
C 108	15	78.9	43986	3	US-09-949-016-17155	Sequence 17155, A	C 181	14.8	77.9	23849	3	US-09-949-016-13813	Sequence 13813, A
C 109	14.8	77.9	247	3	US-09-533-559-7849	Sequence 7849, Ap	C 182	14.8	77.9	24214	3	US-09-949-016-15551	Sequence 15551, A
C 110	14.8	77.9	300	4	US-09-297-648-2037	Sequence 2037, Ap	C 183	14.8	77.9	24735	3	US-09-949-002-716	Sequence 12850, A
C 111	14.8	77.9	363	3	US-09-513-999C-35479	Sequence 35479, A	C 184	14.8	77.9	25589	3	US-09-949-016-12850	Sequence 1716, App
C 112	14.8	77.9	372	3	US-09-513-999C-20504	Sequence 20504, A	C 185	14.8	77.9	26011	3	US-09-949-016-17499	Sequence 17499, A
C 113	14.8	77.9	378	3	US-09-513-999C-21946	Sequence 21946, A	C 186	14.8	77.9	29973	5	US-09-984-429-262	Sequence 262, App
C 114	14.8	77.9	473	3	US-09-513-999C-11142	Sequence 11142, A	C 187	14.8	77.9	43086	3	US-09-949-016-13059	Sequence 13059, A
C 115	14.8	77.9	601	3	US-09-949-016-1141	Sequence 1141, A	C 188	14.8	77.9	43307	3	US-08-949-016-15774	Sequence 15774, A
C 116	14.8	77.9	601	3	US-09-949-016-40853	Sequence 40853, A	C 189	14.8	77.9	64049	3	US-09-949-016-15744	Sequence 17017, A
C 117	14.8	77.9	601	3	US-09-949-016-50658	Sequence 50658, A	C 190	14.8	77.9	68667	3	US-09-949-016-17017	Sequence 17017, A
C 118	14.8	77.9	601	3	US-09-949-016-54267	Sequence 54267, A	C 191	14.8	77.9	70947	3	US-09-949-016-15165	Sequence 15165, A
C 119	14.8	77.9	601	3	US-09-949-016-54331	Sequence 54331, A	C 192	14.8	77.9	78157	3	US-09-949-016-16466	Sequence 16466, A
C 120	14.8	77.9	601	3	US-09-949-016-69064	Sequence 69064, A	C 193	14.8	77.9	78157	3	US-09-949-016-16467	Sequence 16467, A
C 121	14.8	77.9	601	3	US-09-949-016-70530	Sequence 70530, A	C 194	14.8	77.9	78649	3	US-08-949-016-14819	Sequence 14619, A
C 122	14.8	77.9	601	3	US-09-949-016-77508	Sequence 77508, A	C 195	14.8	77.9	78649	3	US-09-949-016-14820	Sequence 14620, A
C 123	14.8	77.9	601	3	US-09-949-016-77509	Sequence 77509, A	C 196	14.8	77.9	78649	3	US-09-949-016-16327	Sequence 16227, A
C 124	14.8	77.9	601	3	US-09-949-016-94629	Sequence 94629, A	C 197	14.8	77.9	78649	3	US-09-949-016-16328	Sequence 14489, A
C 125	14.8	77.9	601	3	US-09-949-016-122013	Sequence 122013, A	C 198	14.8	77.9	83218	3	US-09-949-016-14489	Sequence 14489, A
C 126	14.8	77.9	601	3	US-09-949-016-122014	Sequence 122014, A	C 199	14.8	77.9	83343	3	US-09-949-016-12242	Sequence 12242, A
C 127	14.8	77.9	601	3	US-09-949-016-134694	Sequence 134694, A	C 200	14.8	77.9	83497	3	US-09-949-016-15517	Sequence 15517, A
C 128	14.8	77.9	601	3	US-09-949-016-144401	Sequence 144401, A	C 201	14.8	77.9	85205	3	US-09-949-016-12211	Sequence 12211, A
C 129	14.8	77.9	601	3	US-09-949-016-167556	Sequence 167556, A	C 202	14.8	77.9	87305	3	US-09-949-016-13430	Sequence 13430, A
C 130	14.8	77.9	601	3	US-09-949-016-167663	Sequence 167663, A	C 203	14.8	77.9	90776	3	US-09-949-016-17230	Sequence 17230, A
C 131	14.8	77.9	601	3	US-09-949-016-181131	Sequence 181131, A	C 204	14.8	77.9	92407	3	US-09-596-002-36	Sequence 36, Appl
C 132	14.8	77.9	601	3	US-09-949-016-181132	Sequence 181132, A	C 205	14.8	77.9	93493	3	US-09-949-016-12063	Sequence 12063, A
C 133	14.8	77.9	601	3	US-09-949-016-181133	Sequence 181133, A	C 206	14.8	77.9	95020	3	US-09-949-016-13272	Sequence 13272, A
C 134	14.8	77.9	601	3	US-09-949-002-5993	Sequence 993, App	C 207	14.8	77.9	95318	3	US-09-949-016-11784	Sequence 11784, A
C 135	14.8	77.9	601	3	US-09-949-002-994	Sequence 994, App	C 208	14.8	77.9	95318	3	US-09-949-016-12998	Sequence 13998, A
C 136	14.8	77.9	601	3	US-09-949-002-4238	Sequence 4238, Ap	C 209	14.8	77.9	95561	3	US-09-949-016-12768	Sequence 12768, A
C 137	14.8	77.9	601	3	US-09-949-002-4239	Sequence 4239, Ap	C 210	14.8	77.9	95561	3	US-09-949-016-13306	Sequence 13306, A
C 138	14.8	77.9	662	3	US-09-270-767-28813	Sequence 28813, A	C 211	14.8	77.9	95561	3	US-09-949-016-13307	Sequence 13307, A
C 139	14.8	77.9	676	3	US-09-222-575-41	Sequence 41, Appl	C 212	14.8	77.9	96866	3	US-09-949-016-13768	Sequence 13768, A
C 140	14.8	77.9	676	3	US-09-389-681-41	Sequence 41, Appl	C 213	14.8	77.9	99916	3	US-09-816-095-3	Sequence 3, Appli
C 141	14.8	77.9	676	3	US-09-620-405B-41	Sequence 41, Appl	C 214	14.8	77.9	101786	4	US-09-531-120-199	Sequence 199, App
C 142	14.8	77.9	676	3	US-09-339-338-41	Sequence 41, Appl	C 215	14.8	77.9	105055	3	US-09-949-016-14001	Sequence 14001, A
C 143	14.8	77.9	676	3	US-09-433-826B-41	Sequence 41, Appl	C 216	14.8	77.9	109925	3	US-09-949-016-13210	Sequence 13210, A
C 144	14.8	77.9	676	3	US-09-604-287A-41	Sequence 41, Appl	C 217	14.8	77.9	116177	3	US-09-949-016-16946	Sequence 16946, A
C 145	14.8	77.9	676	3	US-09-285-480-41	Sequence 41, Appl	C 218	14.8	77.9	116373	3	US-09-949-016-12273	Sequence 12273, A
C 146	14.8	77.9	676	3	US-09-834-759-41	Sequence 41, Appl	C 219	14.8	77.9	130724	3	US-09-949-016-13753	Sequence 13753, A
C 147	14.8	77.9	676	3	US-09-530-751A-41	Sequence 41, Appl	C 220	14.8	77.9	131379	3	US-09-949-016-16050	Sequence 16050, A
C 148	14.8	77.9	676	3	US-09-551-621-41	Sequence 41, Appl	C 221	14.8	77.9	133559	3	US-09-949-016-15845	Sequence 15845, A
C 149	14.8	77.9	676	3	US-09-551-621A-41	Sequence 41, Appl	C 222	14.8	77.9	133613	3	US-09-949-016-15824	Sequence 15824, A
C 150	14.8	77.9	676	3	US-10-076-622-41	Sequence 41, Appl	C 223	14.8	77.9	141589	5	US-09-543-679A-2480	Sequence 2480, Ap
C 151	14.8	77.9	676	4	US-10-124-805-41	Sequence 41, Appl	C 224	14.8	77.9	141589	5	US-09-543-679A-2694	Sequence 2694, Ap
C 152	14.8	77.9	707	5	US-09-974-300-5294	Sequence 5294, Ap	C 225	14.8	77.9	141589	5	US-09-543-679A-2719	Sequence 2719, Ap
C 153	14.8	77.9	708	5	US-09-974-300-5134	Sequence 5134, Ap	C 226	14.8	77.9	146982	5	US-09-543-679A-3009	Sequence 3009, Ap
C 154	14.8	77.9	852	3	US-09-248-796A-2002	Sequence 2002, Ap	C 227	14.8	77.9	151828	4	US-09-531-120-197	Sequence 197, App
C 155	14.8	77.9	936	3	US-09-270-767-12940	Sequence 12940, A	C 228	14.8	77.9	161124	3	US-09-949-016-11760	Sequence 1760, A
C 156	14.8	77.9	966	3	US-08-748-506-8	Sequence 8, Appli	C 229	14.8	77.9	176373	3	US-09-128-155-17	Sequence 17, Appl
C 157	14.8	77.9	1001	3	US-09-641-638-388	Sequence 388, App	C 230	14.8	77.9	179904	3	US-09-949-002-577	Sequence 577, App
C 158	14.8	77.9	1001	3	US-09-641-638-390	Sequence 390, App	C 231	14.8	77.9	179905	3	US-09-949-002-705	Sequence 705, App
C 159	14.8	77.9	1001	3	US-10-170-097-388	Sequence 388, App	C 232	14.8	77.9	203475	3	US-09-949-016-14516	Sequence 14516, A
C 160	14.8	77.9	1001	3	US-10-170-097-390	Sequence 390, App	C 233	14.8	77.9	203475	3	US-09-949-016-14517	Sequence 14517, A
C 161	14.8	77.9	1177	2	US-08-249-671A-8	Sequence 8, Appli	C 234	14.8	77.9	203475	3	US-09-949-016-14518	Sequence 14518, A
C 162	14.8	77.9	1236	3	US-09-799-451-867	Sequence 867, App	C 235	14.8	77.9	203475	3	US-09-949-016-14519	Sequence 14519, A
C 163	14.8	77.9	1320	3	US-09-023-655-491	Sequence 491, App	C 236	14.8	77.9	203475	3	US-09-949-016-17226	Sequence 17226, A
C 164	14.8	77.9	2102	3	US-09-023-655-364	Sequence 364, App	C 237	14.8	77.9	203475	3	US-09-949-016-17227	Sequence 17227, A
C 165	14.8	77.9	2313	3	US-09-107-532A-734	Sequence 734, App	C 238	14.8	77.9	203475	3	US-09-949-016-17228	Sequence 17228, A
C 166	14.8	77.9	2457	4	US-10-094-749-438	Sequence 438, App	C 239	14.8	77.9	203475	3	US-09-949-016-17229	Sequence 17229, A
C 167	14.8	77.9	2760	3	US-09-543-681A-86	Sequence 86, Appl	C 240	14.8	77.9	209274	5	US-09-543-679A-3004	Sequence 3004, Ap
C 168	14.8	77.9	3089	3	US-09-016-434-1330	Sequence 1330, Ap	C 241	14.8	77.9	260286	3	US-09-949-016-17037	Sequence 17037, A
C 169	14.8	77.9	3089	4	US-09-880-107-3687	Sequence 3687, Ap	C 242	14.8	77.9	260293	3	US-09-949-016-12106	Sequence 12106, A

243	14.8	77.9	267482	3	US-09-949-002-659	Sequence 659, App	C 316	14.2	74.7	131	3	US-10-150-104-1	Sequence 1, Appli
244	14.8	77.9	267505	3	US-09-949-002-783	Sequence 783, App	317	14.2	74.7	198	5	US-09-974-300-6246	Sequence 6246, Ap
245	14.8	77.9	292500	3	US-10-027-983-11	Sequence 11, Appl	C 318	14.2	74.7	267	3	US-09-313-294A-1237	Sequence 1237, Ap
246	14.8	77.9	462589	3	US-09-949-016-12900	Sequence 12900, A	C 319	14.2	74.7	300	4	US-09-297-648-2319	Sequence 2319, Ap
247	14.8	77.9	476044	3	US-09-949-016-12412	Sequence 12412, A	C 320	14.2	74.7	319	3	US-09-221-017B-458	Sequence 458, App
248	14.8	77.9	611587	4	US-09-531-120-209	Sequence 209, App	C 321	14.2	74.7	360	3	US-09-270-0767-29190	Sequence 29190, A
249	14.8	77.9	678533	3	US-09-949-016-14577	Sequence 14577, A	C 322	14.2	74.7	378	3	US-10-002-623-563	Sequence 563, App
250	14.8	77.9	678533	3	US-09-949-016-14578	Sequence 14578, A	C 323	14.2	74.7	449	3	US-09-513-999C-917	Sequence 917, App
251	14.8	77.9	1664976	3	US-08-916-421B-1	Sequence 1, Appli	C 324	14.2	74.7	476	3	US-09-513-999C-11720	Sequence 11720, A
252	14.8	77.9	1664976	3	US-09-692-570-1	Sequence 1, Appli	C 325	14.2	74.7	480	3	US-09-198-930-206	Sequence 206, App
C 253	14.8	77.9	1830121	3	US-09-557-884-1	Sequence 1, Appli	326	14.2	74.7	480	3	US-09-312-283C-206	Sequence 206, App
C 254	14.8	77.9	1830121	3	US-09-643-990A-1	Sequence 1, Appli	327	14.2	74.7	480	3	US-09-107-532A-420	Sequence 420, App
C 255	14.8	77.9	1830121	3	US-10-158-865-1	Sequence 1, Appli	C 328	14.2	74.7	494	3	US-09-513-999C-892	Sequence 892, App
256	14.4	75.8	47	3	US-09-641-638-1043	Sequence 1043, Ap	C 329	14.2	74.7	549	3	US-09-854-133-515	Sequence 515, App
257	14.4	75.8	312	3	US-10-170-097-1043	Sequence 1043, Ap	C 330	14.2	74.7	573	3	US-09-248-796A-3436	Sequence 3436, Ap
C 258	14.4	75.8	312	3	US-09-248-796A-2555	Sequence 2555, Ap	C 331	14.2	74.7	601	3	US-09-949-016-58296	Sequence 58296, A
C 259	14.4	75.8	356	2	US-08-606-789-15	Sequence 15, Appl	C 332	14.2	74.7	601	3	US-09-949-016-60332	Sequence 60332, A
260	14.4	75.8	356	2	US-09-111-348-15	Sequence 15, Appl	C 333	14.2	74.7	601	3	US-09-949-016-74393	Sequence 74393, A
261	14.4	75.8	437	4	US-09-880-107-2647	Sequence 2647, Ap	C 334	14.2	74.7	601	3	US-09-949-016-76486	Sequence 76486, A
C 262	14.4	75.8	565	3	US-09-257-584-10	Sequence 10, Appl	C 335	14.2	74.7	601	3	US-09-949-016-89613	Sequence 89613, A
C 263	14.4	75.8	601	3	US-09-949-016-44349	Sequence 44349, A	C 336	14.2	74.7	601	3	US-09-949-016-89614	Sequence 89614, A
C 264	14.4	75.8	601	3	US-09-949-016-57961	Sequence 57961, A	C 337	14.2	74.7	601	3	US-09-949-016-112118	Sequence 112118, A
C 265	14.4	75.8	601	3	US-09-949-016-57962	Sequence 57962, A	C 338	14.2	74.7	601	3	US-09-949-016-119242	Sequence 119242, A
266	14.4	75.8	601	3	US-09-949-016-64318	Sequence 64318, A	C 339	14.2	74.7	601	3	US-09-949-016-119373	Sequence 119373, A
C 267	14.4	75.8	601	3	US-09-949-016-64381	Sequence 64381, A	C 340	14.2	74.7	601	3	US-09-949-016-123966	Sequence 123966, A
C 268	14.4	75.8	601	3	US-09-949-016-76382	Sequence 76382, A	C 341	14.2	74.7	601	3	US-09-949-016-126758	Sequence 126758, A
269	14.4	75.8	601	3	US-09-949-016-87508	Sequence 87508, A	C 342	14.2	74.7	601	3	US-09-949-016-126759	Sequence 126759, A
C 270	14.4	75.8	601	3	US-09-949-016-124311	Sequence 124311, A	C 343	14.2	74.7	601	3	US-09-949-016-126807	Sequence 126807, A
C 271	14.4	75.8	601	3	US-09-949-016-144426	Sequence 144426, A	C 344	14.2	74.7	601	3	US-09-949-016-126808	Sequence 126808, A
C 272	14.4	75.8	601	3	US-09-949-016-150280	Sequence 150280, A	C 345	14.2	74.7	601	3	US-09-949-016-126856	Sequence 126856, A
C 273	14.4	75.8	601	3	US-09-949-016-177324	Sequence 177324, A	C 346	14.2	74.7	601	3	US-09-949-016-126857	Sequence 126857, A
C 274	14.4	75.8	801	5	US-09-974-300-8142	Sequence 8142, Ap	C 347	14.2	74.7	601	3	US-09-949-016-133593	Sequence 133593, A
275	14.4	75.8	1001	3	US-09-641-638-389	Sequence 389, App	C 348	14.2	74.7	601	3	US-09-949-016-134391	Sequence 134391, A
C 276	14.4	75.8	1001	3	US-10-170-097-389	Sequence 389, App	C 349	14.2	74.7	601	3	US-09-949-016-134392	Sequence 134392, A
C 277	14.4	75.8	1082	3	US-09-484-970B-109	Sequence 109, App	C 350	14.2	74.7	601	3	US-09-949-016-134440	Sequence 134440, A
C 278	14.4	75.8	1185	3	US-09-248-796A-1925	Sequence 1925, Ap	C 351	14.2	74.7	601	3	US-09-949-016-134441	Sequence 134441, A
C 279	14.4	75.8	1563	3	US-09-377-557-11	Sequence 11, Appl	C 352	14.2	74.7	601	3	US-09-949-016-134489	Sequence 134489, A
C 280	14.4	75.8	2706	2	US-08-630-822A-61	Sequence 61, Appl	C 353	14.2	74.7	601	3	US-09-949-016-134490	Sequence 134490, A
C 281	14.4	75.8	2706	2	US-09-005-069-61	Sequence 61, Appl	C 354	14.2	74.7	601	3	US-09-949-016-152082	Sequence 152082, A
C 282	14.4	75.8	2706	3	US-09-171-156A-20	Sequence 20, Appl	C 355	14.2	74.7	601	3	US-09-949-016-152083	Sequence 152083, A
C 283	14.4	75.8	2706	3	US-09-004-730A-20	Sequence 20, Appl	C 356	14.2	74.7	601	3	US-09-949-016-152084	Sequence 152084, A
C 284	14.4	75.8	2706	3	US-08-981-799A-20	Sequence 20, Appl	C 357	14.2	74.7	601	3	US-09-949-016-152155	Sequence 152155, A
C 285	14.4	75.8	2847	3	US-09-799-451-513	Sequence 513, App	C 358	14.2	74.7	601	3	US-09-949-016-152156	Sequence 152156, A
C 286	14.4	75.8	8640	3	US-09-964-956-24	Sequence 24, Appl	C 359	14.2	74.7	601	3	US-09-949-016-152157	Sequence 152157, A
287	14.4	75.8	8640	3	US-09-964-956-26	Sequence 26, Appl	C 360	14.2	74.7	601	3	US-09-949-016-152228	Sequence 152228, A
C 288	14.4	75.8	18367	3	US-09-949-016-14065	Sequence 14065, A	C 361	14.2	74.7	601	3	US-09-949-016-152229	Sequence 152229, A
C 289	14.4	75.8	21516	3	US-09-949-016-13024	Sequence 13024, A	C 362	14.2	74.7	601	3	US-09-949-016-152230	Sequence 152230, A
290	14.4	75.8	29636	3	US-09-949-016-12189	Sequence 12189, A	C 363	14.2	74.7	601	3	US-09-949-016-152301	Sequence 152301, A
291	14.4	75.8	29637	3	US-09-949-016-17173	Sequence 17173, A	C 364	14.2	74.7	601	3	US-09-949-016-152302	Sequence 152302, A
292	14.4	75.8	29638	3	US-09-949-016-14137	Sequence 14137, A	C 365	14.2	74.7	601	3	US-09-949-016-152303	Sequence 152303, A
293	14.4	75.8	33478	3	US-09-949-016-13615	Sequence 13615, A	C 366	14.2	74.7	601	3	US-09-949-016-154785	Sequence 154785, A
C 294	14.4	75.8	36047	3	US-09-949-016-13966	Sequence 13966, A	C 367	14.2	74.7	601	3	US-09-949-016-159005	Sequence 159005, A
C 295	14.4	75.8	47968	3	US-09-949-016-15240	Sequence 15240, A	C 368	14.2	74.7	601	3	US-09-949-016-159387	Sequence 159387, A
296	14.4	75.8	68173	3	US-09-949-016-14046	Sequence 14046, A	C 369	14.2	74.7	601	3	US-09-949-016-159388	Sequence 159388, A
297	14.4	75.8	69909	3	US-09-949-016-13423	Sequence 13423, A	C 370	14.2	74.7	601	3	US-09-949-016-159389	Sequence 159389, A
298	14.4	75.8	77535	3	US-09-949-016-14279	Sequence 14279, A	C 371	14.2	74.7	601	3	US-09-949-016-159460	Sequence 159460, A
299	14.4	75.8	77535	3	US-09-949-016-14280	Sequence 14280, A	C 372	14.2	74.7	601	3	US-09-949-016-159461	Sequence 159461, A
C 300	14.4	75.8	77535	3	US-09-949-016-14281	Sequence 14281, A	C 373	14.2	74.7	601	3	US-09-949-016-159462	Sequence 159462, A
C 301	14.4	75.8	84252	3	US-09-949-016-17315	Sequence 17315, A	C 374	14.2	74.7	601	3	US-09-949-016-159533	Sequence 159533, A
C 302	14.4	75.8	96048	3	US-09-949-016-15966	Sequence 15966, A	C 375	14.2	74.7	601	3	US-09-949-016-159534	Sequence 159534, A
C 303	14.4	75.8	133613	3	US-09-949-016-15824	Sequence 15824, A	C 376	14.2	74.7	601	3	US-09-949-016-159535	Sequence 159535, A
C 304	14.4	75.8	144158	3	US-09-949-016-11755	Sequence 11755, A	C 377	14.2	74.7	601	3	US-09-949-016-154785	Sequence 154785, A
C 305	14.4	75.8	144158	3	US-09-949-016-12936	Sequence 12936, A	C 378	14.2	74.7	601	3	US-09-949-016-159607	Sequence 159607, A
C 306	14.4	75.8	145241	3	US-09-949-016-17394	Sequence 17394, A	C 379	14.2	74.7	601	3	US-09-949-016-159608	Sequence 159608, A
307	14.4	75.8	145241	3	US-09-949-016-17395	Sequence 17395, A	C 380	14.2	74.7	601	3	US-09-949-016-161881	Sequence 161881, A
308	14.4	75.8	176006	3	US-09-949-016-16804	Sequence 16804, A	C 381	14.2	74.7	601	3	US-09-949-016-166114	Sequence 166114, A
309	14.4	75.8	182700	3	US-09-949-016-13266	Sequence 13266, A	C 382	14.2	74.7	601	3	US-09-949-016-171581	Sequence 171581, A
310	14.4	75.8	192700	3	US-09-949-016-11820	Sequence 11820, A	C 383	14.2	74.7	601	3	US-09-949-016-179311	Sequence 179311, A
311	14.4	75.8	192704	3	US-09-949-016-17182	Sequence 17182, A	C 384	14.2	74.7	601	3	US-09-949-016-179312	Sequence 179312, A
312	14.4	75.8	233375	3	US-09-949-016-12849	Sequence 12849, A	C 385	14.2	74.7	601	3	US-09-949-016-186673	Sequence 186673, A
313	14.4	75.8	283538	3	US-09-949-016-13506	Sequence 13506, A	C 386	14.2	74.7	601	3	US-09-949-016-186674	Sequence 186674, A
314	14.4	75.8	670689	3	US-09-949-016-12505	Sequence 12505, A	C 387	14.2	74.7	601	3	US-09-949-016-186675	Sequence 186675, A
315	14.4	75.8	670690	3	US-09-949-016-14207	Sequence 14207, A	C 388	14.2	74.7	601	3	US-09-949-016-201189	Sequence 201189, A

389	14.2	74.7	601	3	US-09-949-016-205561	Sequence 205561,	C 462	14.2	74.7	4680	2	US-08-254-358-1	Sequence 1, Appli
C 390	14.2	74.7	601	3	US-09-949-016-206956	Sequence 206956,	C 463	14.2	74.7	4680	2	US-08-475-391-1	Sequence 1, Appli
C 391	14.2	74.7	601	3	US-09-949-016-206957	Sequence 206957,	C 464	14.2	74.7	4680	2	US-08-709-609-1	Sequence 1, Appli
C 392	14.2	74.7	601	3	US-09-949-016-206958	Sequence 206958,	C 465	14.2	74.7	4680	7	PCT-US95-07178-1	Sequence 1, Appli
C 393	14.2	74.7	601	3	US-09-949-002-1904	Sequence 1904, Ap	C 466	14.2	74.7	4681	3	US-09-807-802A-18	Sequence 18, Appl
C 394	14.2	74.7	601	3	US-09-949-002-1905	Sequence 1905, Ap	C 467	14.2	74.7	4800	3	US-09-949-016-15055	Sequence 15055, A
C 395	14.2	74.7	601	3	US-09-949-002-5851	Sequence 5851, Ap	C 468	14.2	74.7	4833	3	US-09-949-016-14282	Sequence 14282, A
C 396	14.2	74.7	601	3	US-09-949-002-5852	Sequence 5852, Ap	C 469	14.2	74.7	5761	2	US-07-749-001-2	Sequence 2, Appli
C 397	14.2	74.7	611	3	US-09-385-982-378	Sequence 378, App	C 470	14.2	74.7	5761	2	US-08-154-198-2	Sequence 2, Appli
C 398	14.2	74.7	683	5	US-09-974-300-507	Sequence 507, App	C 471	14.2	74.7	5761	2	US-08-463-335-2	Sequence 2, Appli
C 399	14.2	74.7	700	3	US-09-735-271-266	Sequence 266, App	C 472	14.2	74.7	5761	2	US-08-464-023A-2	Sequence 2, Appli
C 400	14.2	74.7	708	3	US-09-533-559-7504	Sequence 7504, Ap	C 473	14.2	74.7	5922	3	US-08-961-527-267	Sequence 267, App
C 401	14.2	74.7	715	3	US-08-998-416-522	Sequence 522, App	C 474	14.2	74.7	6708	3	US-09-949-016-12933	Sequence 12933, A
C 402	14.2	74.7	728	3	US-09-988-863A-4	Sequence 4, Appli	C 475	14.2	74.7	7557	3	US-09-770-315-3	Sequence 3, Appli
C 403	14.2	74.7	749	3	US-09-270-767-13258	Sequence 13258, A	C 476	14.2	74.7	8151	3	US-09-438-268-2	Sequence 2, Appli
C 404	14.2	74.7	778	4	US-09-297-648-3659	Sequence 3659, A	C 477	14.2	74.7	8201	4	US-09-958-617A-17	Sequence 17, Appl
C 405	14.2	74.7	800	3	US-09-949-016-3313	Sequence 3313, Ap	C 478	14.2	74.7	8698	3	US-09-770-315-2	Sequence 2, Appli
C 406	14.2	74.7	811	3	US-09-949-016-3371	Sequence 3371, Ap	C 479	14.2	74.7	8746	3	US-09-949-016-16984	Sequence 16984, A
C 407	14.2	74.7	820	3	US-09-949-016-2540	Sequence 2540, Ap	C 480	14.2	74.7	9192	3	US-09-949-016-14830	Sequence 14830, A
C 408	14.2	74.7	896	3	US-09-188-930-36	Sequence 36, Appl	C 481	14.2	74.7	10207	2	US-08-920-812-2	Sequence 2, Appli
C 409	14.2	74.7	896	3	US-09-312-283C-36	Sequence 36, Appl	C 482	14.2	74.7	10207	2	US-08-920-827-2	Sequence 2, Appli
C 410	14.2	74.7	903	5	US-09-974-300-1703	Sequence 1703, Ap	C 483	14.2	74.7	10207	2	US-08-921-177-2	Sequence 2, Appli
C 411	14.2	74.7	975	3	US-09-489-039A-1422	Sequence 1422, Ap	C 484	14.2	74.7	10207	2	US-08-362-577C-2	Sequence 2, Appli
C 412	14.2	74.7	983	3	US-09-270-767-11267	Sequence 11267, A	C 485	14.2	74.7	10207	2	US-08-920-828-2	Sequence 2, Appli
C 413	14.2	74.7	990	3	US-09-134-000C-1367	Sequence 1367, Ap	C 486	14.2	74.7	11728	3	US-09-949-016-14245	Sequence 14245, A
C 414	14.2	74.7	1049	3	US-09-799-451-892	Sequence 892, App	C 487	14.2	74.7	12139	3	US-09-949-016-14218	Sequence 14218, A
C 415	14.2	74.7	1107	5	US-09-974-300-1582	Sequence 1582, Ap	C 488	14.2	74.7	12191	3	US-09-949-016-16237	Sequence 16237, A
C 416	14.2	74.7	1215	3	US-09-328-352-2728	Sequence 2728, Ap	C 489	14.2	74.7	12423	3	US-09-949-016-14959	Sequence 14959, A
C 417	14.2	74.7	1317	3	US-09-350-756-5	Sequence 5, Appli	C 490	14.2	74.7	12616	3	US-09-949-016-14217	Sequence 14217, A
C 418	14.2	74.7	1323	3	US-09-583-110-1395	Sequence 1395, Ap	C 491	14.2	74.7	16995	3	US-08-961-527-82	Sequence 82, Appl
C 419	14.2	74.7	1338	3	US-09-107-433-915	Sequence 915, App	C 492	14.2	74.7	18026	3	US-09-949-016-13309	Sequence 13309, A
C 420	14.2	74.7	1352	3	US-09-721-1307-15	Sequence 15, Appl	C 493	14.2	74.7	19062	5	US-09-543-679A-2847	Sequence 2847, Ap
C 421	14.2	74.7	1386	3	US-09-134-000C-2146	Sequence 2146, Ap	C 494	14.2	74.7	19389	3	US-09-949-016-15113	Sequence 15113, A
C 422	14.2	74.7	1446	3	US-09-252-991A-16240	Sequence 16240, A	C 495	14.2	74.7	20079	3	US-09-949-016-16060	Sequence 16060, A
C 423	14.2	74.7	1500	3	US-09-252-991A-16351	Sequence 16351, A	C 496	14.2	74.7	25992	3	US-09-949-016-17308	Sequence 17308, A
C 424	14.2	74.7	1842	3	US-09-016-434-46	Sequence 46, Appl	C 497	14.2	74.7	26000	3	US-09-843-376-10	Sequence 10, Appl
C 425	14.2	74.7	1887	4	US-10-099-322-1	Sequence 1, Appli	C 498	14.2	74.7	30417	5	US-09-543-679A-2849	Sequence 2849, Ap
C 426	14.2	74.7	1887	4	US-10-044-564-1	Sequence 1, Appli	C 499	14.2	74.7	31618	3	US-09-949-016-14759	Sequence 14759, A
C 427	14.2	74.7	1973	10	5256643-1	Patent No. 5256643	C 500	14.2	74.7	36156	3	US-09-949-016-12128	Sequence 12128, A
C 428	14.2	74.7	2033	3	US-09-023-655-1266	Sequence 1266, Ap	C 501	14.2	74.7	36156	3	US-09-949-016-13261	Sequence 13261, A
C 429	14.2	74.7	2033	3	US-09-949-016-5242	Sequence 5242, Ap	C 502	14.2	74.7	37925	5	US-09-984-429-701	Sequence 701, App
C 430	14.2	74.7	2160	3	US-09-107-532A-2338	Sequence 2338, Ap	C 503	14.2	74.7	47925	3	US-09-949-016-15829	Sequence 15829, A
C 431	14.2	74.7	2184	4	US-10-094-749-322	Sequence 322, App	C 504	14.2	74.7	45432	3	US-09-949-002-719	Sequence 719, App
C 432	14.2	74.7	2251	3	US-10-104-047-916	Sequence 916, App	C 505	14.2	74.7	47084	3	US-09-949-016-13380	Sequence 13380, A
C 433	14.2	74.7	2361	3	US-09-107-433-1290	Sequence 1290, Ap	C 506	14.2	74.7	47110	3	US-09-949-016-17609	Sequence 17609, A
C 434	14.2	74.7	2370	4	US-10-094-749-2	Sequence 2, Appli	C 507	14.2	74.7	51719	3	US-09-918-686-2	Sequence 2, Appli
C 435	14.2	74.7	2427	3	US-09-543-681A-823	Sequence 823, App	C 508	14.2	74.7	54180	3	US-09-949-016-14894	Sequence 14894, A
C 436	14.2	74.7	2479	3	US-09-949-016-3088	Sequence 3088, Ap	C 509	14.2	74.7	61158	3	US-09-949-016-15041	Sequence 15041, A
C 437	14.2	74.7	2519	5	US-10-770-127-202	Sequence 202, App	C 510	14.2	74.7	62311	3	US-09-949-016-14582	Sequence 14582, A
C 438	14.2	74.7	2520	3	US-09-897-427A-3	Sequence 3, Appli	C 511	14.2	74.7	70323	3	US-09-949-016-17594	Sequence 17594, A
C 439	14.2	74.7	2581	3	US-09-363-708-1	Sequence 1, Appli	C 512	14.2	74.7	75431	3	US-09-949-016-15122	Sequence 15122, A
C 440	14.2	74.7	2581	3	US-09-083-587-1	Sequence 1, Appli	C 513	14.2	74.7	76553	3	US-09-949-016-13432	Sequence 13432, A
C 441	14.2	74.7	2604	3	US-09-248-796A-5104	Sequence 5104, Ap	C 514	14.2	74.7	78846	3	US-09-949-016-12396	Sequence 12396, A
C 442	14.2	74.7	2667	3	US-09-614-221A-335	Sequence 335, App	C 515	14.2	74.7	78846	3	US-09-949-016-12791	Sequence 12791, A
C 443	14.2	74.7	2675	2	US-07-749-001-4	Sequence 4, Appli	C 516	14.2	74.7	78846	3	US-09-949-016-12792	Sequence 12792, A
C 444	14.2	74.7	2675	2	US-08-154-198-4	Sequence 4, Appli	C 517	14.2	74.7	78846	3	US-09-949-016-12793	Sequence 12793, A
C 445	14.2	74.7	2675	2	US-08-463-335-4	Sequence 4, Appli	C 518	14.2	74.7	78850	3	US-09-949-016-16013	Sequence 16013, A
C 446	14.2	74.7	2675	2	US-08-464-023A-4	Sequence 4, Appli	C 519	14.2	74.7	78850	3	US-09-949-016-16014	Sequence 16014, A
C 447	14.2	74.7	2708	3	US-09-949-016-1191	Sequence 1191, Ap	C 520	14.2	74.7	78850	3	US-09-949-016-16015	Sequence 16015, A
C 448	14.2	74.7	2718	3	US-09-248-796A-2230	Sequence 2230, Ap	C 521	14.2	74.7	78850	3	US-09-949-016-16016	Sequence 16016, A
C 449	14.2	74.7	2767	3	US-09-306-595C-4	Sequence 4, Appli	C 522	14.2	74.7	78850	3	US-09-949-016-16201	Sequence 16201, A
C 450	14.2	74.7	2767	3	US-09-925-388-4	Sequence 4, Appli	C 523	14.2	74.7	78850	3	US-09-949-016-16202	Sequence 16202, A
C 451	14.2	74.7	2866	3	US-10-431-846-4	Sequence 4, Appli	C 524	14.2	74.7	78850	3	US-09-949-016-16203	Sequence 16203, A
C 452	14.2	74.7	2893	3	US-10-104-047-1267	Sequence 1267, Ap	C 525	14.2	74.7	78850	3	US-09-949-016-16204	Sequence 16204, A
C 453	14.2	74.7	2998	4	US-10-094-749-651	Sequence 651, App	C 526	14.2	74.7	79578	3	US-09-949-016-16339	Sequence 16339, A
C 454	14.2	74.7	3138	3	US-10-104-047-1676	Sequence 1676, Ap	C 527	14.2	74.7	79595	3	US-09-949-016-15318	Sequence 15318, A
C 455	14.2	74.7	4072	3	US-09-770-315-4	Sequence 4, Appli	C 528	14.2	74.7	80490	3	US-09-949-002-733	Sequence 733, App
C 456	14.2	74.7	4407	3	US-09-949-016-1690	Sequence 1690, Ap	C 529	14.2	74.7	82618	3	US-09-949-002-696	Sequence 696, App
C 457	14.2	74.7	4675	3	US-09-782-378A-1	Sequence 1, Appli	C 530	14.2	74.7	82618	3	US-09-949-002-579	Sequence 579, App
C 458	14.2	74.7	4675	3	US-09-782-378A-2	Sequence 2, Appli	C 531	14.2	74.7	85122	3	US-09-949-016-14693	Sequence 14693, A
C 459	14.2	74.7	4675	4	US-10-111-708-1	Sequence 1, Appli	C 532	14.2	74.7	87594	3	US-09-949-016-12135	Sequence 12135, A
C 460	14.2	74.7	4679	3	US-10-038-972A-12	Sequence 12, Appl	C 533	14.2	74.7	87611	3	US-09-949-016-16139	Sequence 16139, A
C 461	14.2	74.7	4679	4	US-09-717-789C-25	Sequence 25, Appl	C 534	14.2	74.7	92119	3	US-09-918-686-1	Sequence 1, Appli

C 535	14.2	74.7	94142	3	US-09-949-016-16553	Sequence 16553, A	608	13.8	72.6	330	3	US-10-017-754-1379	Sequence 1379, Ap
C 536	14.2	74.7	94618	4	US-09-531-120-191	Sequence 191, App	609	13.8	72.6	330	3	US-09-651-563-1379	Sequence 1379, Ap
C 537	14.2	74.7	96922	4	US-09-949-016-17061	Sequence 17061, A	C 610	13.8	72.6	342	3	US-08-134-000C-2932	Sequence 2932, Ap
C 538	14.2	74.7	101786	4	US-09-531-120-199	Sequence 199, App	C 611	13.8	72.6	346	2	US-08-967-101-33	Sequence 33, Appl
C 539	14.2	74.7	105919	4	US-09-949-016-11769	Sequence 11769, A	C 612	13.8	72.6	346	2	US-08-592-541-33	Sequence 33, Appl
C 540	14.2	74.7	110096	4	US-09-880-107-1542	Sequence 1542, Ap	C 613	13.8	72.6	346	2	US-09-124-698-33	Sequence 33, Appl
C 541	14.2	74.7	116652	4	US-09-949-016-13413	Sequence 13413, A	C 614	13.8	72.6	346	3	US-09-127-480-33	Sequence 33, Appl
C 542	14.2	74.7	119214	3	US-09-949-016-12507	Sequence 12507, A	C 615	13.8	72.6	346	3	US-08-496-841C-33	Sequence 33, Appl
C 543	14.2	74.7	120609	3	US-09-949-016-13915	Sequence 13915, A	C 616	13.8	72.6	346	3	US-08-124-523-33	Sequence 33, Appl
C 544	14.2	74.7	128516	3	US-09-949-016-13501	Sequence 13501, A	C 617	13.8	72.6	346	3	US-08-636-796A-33	Sequence 33, Appl
C 545	14.2	74.7	129778	3	US-09-949-016-12191	Sequence 12191, A	C 618	13.8	72.6	346	3	US-08-431-048F-33	Sequence 33, Appl
C 546	14.2	74.7	129778	3	US-09-949-016-17075	Sequence 17075, A	C 619	13.8	72.6	346	4	US-09-513-999C-13247	Sequence 13247, A
C 547	14.2	74.7	133157	3	US-09-949-016-12541	Sequence 12541, A	C 620	13.8	72.6	353	3	US-09-513-999C-12568	Sequence 12568, A
C 548	14.2	74.7	134987	3	US-09-949-016-15348	Sequence 15348, A	C 621	13.8	72.6	381	5	US-09-974-300-1435	Sequence 1435, Ap
C 549	14.2	74.7	134987	3	US-09-949-016-15349	Sequence 15349, A	C 622	13.8	72.6	399	3	US-08-854-133-701	Sequence 701, App
C 550	14.2	74.7	134987	3	US-09-949-016-15350	Sequence 15350, A	C 623	13.8	72.6	429	3	US-08-558-818-6	Sequence 181, Appl
C 551	14.2	74.7	134987	3	US-09-949-016-15507	Sequence 15507, A	C 624	13.8	72.6	459	3	US-08-974-469A-6	Sequence 16276, A
C 552	14.2	74.7	134987	3	US-09-949-016-15508	Sequence 15508, A	C 625	13.8	72.6	462	3	US-09-401-064-164	Sequence 164, App
C 553	14.2	74.7	135667	3	US-09-949-016-15509	Sequence 15509, A	C 626	13.8	72.6	468	3	US-08-502-535B-1	Sequence 12568, A
C 554	14.2	74.7	139552	3	US-09-949-016-15300	Sequence 15300, A	C 627	13.8	72.6	471	2	US-08-908-005A-1	Sequence 1, Appl
C 555	14.2	74.7	139552	3	US-09-949-016-15300	Sequence 15300, A	C 628	13.8	72.6	471	2	US-08-908-005A-1	Sequence 1, Appl
C 556	14.2	74.7	147894	3	US-09-949-016-15236	Sequence 15236, A	C 629	13.8	72.6	471	3	US-08-558-818-6	Sequence 6, Appl
C 557	14.2	74.7	147894	3	US-09-949-016-12751	Sequence 12751, A	C 630	13.8	72.6	471	3	US-08-974-469A-6	Sequence 6, Appl
C 558	14.2	74.7	150423	3	US-09-949-016-17057	Sequence 17057, A	C 631	13.8	72.6	471	3	US-08-832-180-3	Sequence 3, Appl
C 559	14.2	74.7	155617	3	US-09-949-016-16191	Sequence 16191, A	C 632	13.8	72.6	471	3	US-08-832-180-3	Sequence 10, Appl
C 560	14.2	74.7	163664	3	US-09-949-016-12545	Sequence 12545, A	C 633	13.8	72.6	471	3	US-09-253-523-1	Sequence 1, Appl
C 561	14.2	74.7	163664	3	US-09-949-016-12545	Sequence 12545, A	C 634	13.8	72.6	471	3	US-09-253-523-1	Sequence 1, Appl
C 562	14.2	74.7	174029	3	US-09-949-016-12610	Sequence 12610, A	C 635	13.8	72.6	471	3	US-09-819-902-10	Sequence 10, Appl
C 563	14.2	74.7	174030	3	US-09-949-016-13880	Sequence 13880, A	C 636	13.8	72.6	471	3	US-09-752-510-10	Sequence 10, Appl
C 564	14.2	74.7	189560	3	US-09-949-016-17202	Sequence 17202, A	C 637	13.8	72.6	471	3	US-08-982-285-22	Sequence 22, Appl
C 565	14.2	74.7	198632	3	US-09-949-016-12781	Sequence 12781, A	C 638	13.8	72.6	471	3	US-08-982-285-23	Sequence 23, Appl
C 566	14.2	74.7	198637	3	US-09-949-016-12781	Sequence 12781, A	C 639	13.8	72.6	471	3	US-08-982-285-28	Sequence 28, Appl
C 567	14.2	74.7	209631	3	US-09-949-002-574	Sequence 574, App	C 640	13.8	72.6	471	3	US-08-711-899-6	Sequence 6, Appl
C 568	14.2	74.7	209632	3	US-09-949-002-802	Sequence 802, App	C 641	13.8	72.6	471	3	US-10-100-057-18	Sequence 18, Appl
C 569	14.2	74.7	219964	3	US-09-949-016-15086	Sequence 15086, A	C 642	13.8	72.6	474	3	US-10-105-080-2	Sequence 2, Appl
C 570	14.2	74.7	234884	3	US-09-949-016-16420	Sequence 16420, A	C 643	13.8	72.6	474	3	US-10-105-080-9	Sequence 9, Appl
C 571	14.2	74.7	237510	3	US-09-949-016-14273	Sequence 14273, A	C 644	13.8	72.6	513	3	US-09-710-279-1925	Sequence 1925, Ap
C 572	14.2	74.7	238815	3	US-09-949-016-16274	Sequence 16274, A	C 645	13.8	72.6	541	3	US-08-679-409-39	Sequence 39, Appl
C 573	14.2	74.7	242826	3	US-09-949-016-15497	Sequence 15497, A	C 646	13.8	72.6	570	3	US-08-982-285-29	Sequence 29, Appl
C 574	14.2	74.7	256171	3	US-09-949-016-12822	Sequence 12822, A	C 647	13.8	72.6	579	3	US-10-105-080-7	Sequence 7, Appl
C 575	14.2	74.7	256176	3	US-09-949-016-15524	Sequence 15524, A	C 648	13.8	72.6	601	3	US-09-949-016-20143	Sequence 20143, A
C 576	14.2	74.7	256775	3	US-09-949-016-16435	Sequence 16435, A	C 649	13.8	72.6	601	3	US-09-949-016-23899	Sequence 23899, A
C 577	14.2	74.7	263693	3	US-09-949-016-12386	Sequence 12386, A	C 650	13.8	72.6	601	3	US-09-949-016-23900	Sequence 23900, A
C 578	14.2	74.7	263694	3	US-09-949-016-16915	Sequence 16915, A	C 651	13.8	72.6	601	3	US-09-949-016-23901	Sequence 23901, A
C 579	14.2	74.7	271134	3	US-08-949-016-12705	Sequence 12705, A	C 652	13.8	72.6	601	3	US-08-949-016-25725	Sequence 25725, A
C 580	14.2	74.7	275110	3	US-09-949-016-12706	Sequence 12706, A	C 653	13.8	72.6	601	3	US-09-949-016-31201	Sequence 31201, A
C 581	14.2	74.7	275110	3	US-09-949-016-16070	Sequence 16070, A	C 654	13.8	72.6	601	3	US-09-949-016-31457	Sequence 31457, A
C 582	14.2	74.7	298336	3	US-09-949-016-16600	Sequence 16600, A	C 655	13.8	72.6	601	3	US-09-949-016-31458	Sequence 31458, A
C 583	14.2	74.7	301828	3	US-09-949-016-13969	Sequence 13969, A	C 656	13.8	72.6	601	3	US-09-949-016-31459	Sequence 31459, A
C 584	14.2	74.7	305491	3	US-09-949-016-17550	Sequence 17550, A	C 657	13.8	72.6	601	3	US-09-949-016-33735	Sequence 33735, A
C 585	14.2	74.7	317366	3	US-09-949-016-16001	Sequence 16001, A	C 658	13.8	72.6	601	3	US-08-949-016-34165	Sequence 34165, A
C 586	14.2	74.7	346112	3	US-09-949-016-13165	Sequence 13165, A	C 659	13.8	72.6	601	3	US-09-949-016-34576	Sequence 34576, A
C 587	14	73.7	462	3	US-09-270-767-2290	Sequence 2290, Ap	C 660	13.8	72.6	601	3	US-09-949-016-38553	Sequence 38553, A
C 588	14	73.7	462	3	US-09-270-767-17572	Sequence 17572, A	C 661	13.8	72.6	601	3	US-09-949-016-42831	Sequence 42831, A
C 589	14	73.7	497	3	US-09-328-475C-220	Sequence 220, App	C 662	13.8	72.6	601	3	US-09-949-016-45955	Sequence 45955, A
C 590	14	73.7	688	3	US-09-270-767-4835	Sequence 4835, Ap	C 663	13.8	72.6	601	3	US-09-949-016-46247	Sequence 46247, A
C 591	14	73.7	688	3	US-09-270-767-20117	Sequence 20117, A	C 664	13.8	72.6	601	3	US-08-949-016-46248	Sequence 46248, A
C 592	14	73.7	787	3	US-09-328-475C-336	Sequence 336, App	C 665	13.8	72.6	601	3	US-09-949-016-46249	Sequence 46249, A
C 593	14	73.7	831	3	US-09-328-475C-221	Sequence 221, App	C 666	13.8	72.6	601	3	US-09-949-016-48443	Sequence 48443, A
C 594	14	73.7	1236	3	US-09-540-236-1629	Sequence 1629, Ap	C 667	13.8	72.6	601	3	US-09-949-016-48444	Sequence 48444, A
C 595	14	73.7	1236	3	US-09-268-140-6	Sequence 6, Appl	C 668	13.8	72.6	601	3	US-09-949-016-48445	Sequence 48445, A
C 596	14	73.7	99629	3	US-09-596-002-37	Sequence 37, Appl	C 669	13.8	72.6	601	3	US-09-949-016-48716	Sequence 48716, A
C 597	13.8	72.6	187	3	US-08-513-999C-23586	Sequence 23586, A	C 670	13.8	72.6	601	3	US-09-949-016-48717	Sequence 48717, A
C 598	13.8	72.6	249	3	US-09-134-001C-114	Sequence 114, App	C 671	13.8	72.6	601	3	US-09-949-016-52063	Sequence 52063, A
C 599	13.8	72.6	265	3	US-09-513-999C-34761	Sequence 34761, A	C 672	13.8	72.6	601	3	US-09-949-016-54430	Sequence 54430, A
C 600	13.8	72.6	282	5	US-08-974-300-3743	Sequence 3743, Ap	C 673	13.8	72.6	601	3	US-09-949-016-55106	Sequence 55106, A
C 601	13.8	72.6	294	3	US-09-621-976-3324	Sequence 3324, Ap	C 674	13.8	72.6	601	3	US-09-949-016-55438	Sequence 55438, A
C 602	13.8	72.6	305	3	US-09-313-294A-6618	Sequence 6618, Ap	C 675	13.8	72.6	601	3	US-09-949-016-55636	Sequence 55636, A
C 603	13.8	72.6	330	3	US-09-702-705-1379	Sequence 1379, Ap	C 676	13.8	72.6	601	3	US-08-949-016-57018	Sequence 57018, A
C 604	13.8	72.6	330	3	US-09-736-457-1379	Sequence 1379, Ap	C 677	13.8	72.6	601	3	US-09-949-016-58037	Sequence 58037, A
C 605	13.8	72.6	330	3	US-09-614-124B-1379	Sequence 1379, Ap	C 678	13.8	72.6	601	3	US-09-949-016-58503	Sequence 58503, A
C 606	13.8	72.6	330	3	US-09-671-325-1379	Sequence 1379, Ap	C 679	13.8	72.6	601	3	US-09-949-016-61009	Sequence 61009, A
C 607	13.8	72.6	330	3	US-09-658-824-1379	Sequence 1379, Ap	C 680	13.8	72.6	601	3	US-09-949-016-66355	Sequence 66355, A

681	13.8	72.6	601	3	US-09-949-016-68074	Sequence 58074, A	754	13.8	72.6	905	3	US-09-774-639-85	Sequence 85, Appl
c 682	13.8	72.6	601	3	US-09-949-016-77543	Sequence 77543, A	c 755	13.8	72.6	915	3	US-09-248-796A-1167	Sequence 1167, Ap
c 683	13.8	72.6	601	3	US-09-949-016-93849	Sequence 93849, A	c 756	13.8	72.6	915	3	US-09-248-796A-4782	Sequence 4782, Ap
684	13.8	72.6	601	3	US-09-949-016-108134	Sequence 108134,	c 757	13.8	72.6	930	5	US-10-770-127-57	Sequence 57, Appl
685	13.8	72.6	601	3	US-09-949-016-108236	Sequence 108236,	c 758	13.8	72.6	930	5	US-10-770-127-61	Sequence 61, Appl
686	13.8	72.6	601	3	US-09-949-016-108237	Sequence 108237,	c 759	13.8	72.6	930	5	US-10-770-127-63	Sequence 63, Appl
687	13.8	72.6	601	3	US-09-949-016-108236	Sequence 108236,	c 760	13.8	72.6	930	5	US-10-770-127-182	Sequence 182, App
688	13.8	72.6	601	3	US-09-949-016-108297	Sequence 108297,	c 761	13.8	72.6	930	5	US-10-770-127-190	Sequence 190, App
689	13.8	72.6	601	3	US-09-949-016-108326	Sequence 108326,	c 762	13.8	72.6	930	5	US-10-770-127-192	Sequence 192, App
690	13.8	72.6	601	3	US-09-949-016-108327	Sequence 108327,	c 763	13.8	72.6	960	5	US-10-770-127-52	Sequence 52, Appl
691	13.8	72.6	601	3	US-09-949-016-108328	Sequence 108328,	c 764	13.8	72.6	999	3	US-09-248-796A-5959	Sequence 5959, Ap
c 692	13.8	72.6	601	3	US-09-949-016-111665	Sequence 111665,	c 765	13.8	72.6	1083	3	US-09-134-001C-1520	Sequence 1520, Ap
c 693	13.8	72.6	601	3	US-09-949-016-111813	Sequence 111813,	c 766	13.8	72.6	1119	4	US-09-605-703B-531	Sequence 531, App
c 694	13.8	72.6	601	3	US-09-949-016-111959	Sequence 111959,	c 767	13.8	72.6	1172	3	US-09-016-434-1322	Sequence 1322, Ap
c 695	13.8	72.6	601	3	US-09-949-016-112104	Sequence 112104,	c 768	13.8	72.6	1172	3	US-09-023-655-1320	Sequence 1320, Ap
c 696	13.8	72.6	601	3	US-09-949-016-113383	Sequence 113383,	c 769	13.8	72.6	1172	3	US-09-917-254-44	Sequence 44, Appl
c 697	13.8	72.6	601	3	US-09-949-016-127123	Sequence 127123,	c 770	13.8	72.6	1172	5	US-09-913-171A-10	Sequence 10, Appl
c 698	13.8	72.6	601	3	US-09-949-016-127124	Sequence 127124,	c 771	13.8	72.6	1174	2	US-09-663-600A-143	Sequence 143, App
c 699	13.8	72.6	601	3	US-09-949-016-127125	Sequence 127125,	c 772	13.8	72.6	1185	2	US-08-493-638-5	Sequence 5, Appl
c 700	13.8	72.6	601	3	US-09-949-016-133650	Sequence 133650,	c 773	13.8	72.6	1245	3	US-09-614-221A-469	Sequence 469, App
c 701	13.8	72.6	601	3	US-09-949-016-133650	Sequence 133650,	c 774	13.8	72.6	1259	3	US-09-270-767-31484	Sequence 31484, A
702	13.8	72.6	601	3	US-09-949-016-137317	Sequence 137317,	c 775	13.8	72.6	1261	3	US-08-858-207A-151	Sequence 151, App
703	13.8	72.6	601	3	US-09-949-016-139254	Sequence 139254,	c 776	13.8	72.6	1312	3	US-08-961-083-69	Sequence 69, Appl
704	13.8	72.6	601	3	US-09-949-016-139255	Sequence 139255,	c 777	13.8	72.6	1312	3	US-09-536-784-69	Sequence 69, Appl
c 705	13.8	72.6	601	3	US-09-949-016-139256	Sequence 139256,	c 778	13.8	72.6	1312	3	US-09-765-271-69	Sequence 69, Appl
c 706	13.8	72.6	601	3	US-09-949-016-144764	Sequence 144764,	c 779	13.8	72.6	1312	3	US-09-765-272A-69	Sequence 69, Appl
707	13.8	72.6	601	3	US-09-949-016-144765	Sequence 144765,	c 780	13.8	72.6	1410	3	US-09-328-352-2150	Sequence 2150, Ap
708	13.8	72.6	601	3	US-09-949-016-146926	Sequence 146926,	c 781	13.8	72.6	1424	3	US-09-620-405B-464	Sequence 464, App
709	13.8	72.6	601	3	US-09-949-016-146927	Sequence 146927,	c 782	13.8	72.6	1424	3	US-09-433-826B-464	Sequence 464, App
c 710	13.8	72.6	601	3	US-09-949-016-147316	Sequence 147316,	c 783	13.8	72.6	1424	3	US-09-604-287A-464	Sequence 464, App
c 711	13.8	72.6	601	3	US-09-949-016-149205	Sequence 149205,	c 784	13.8	72.6	1424	3	US-09-834-759-464	Sequence 464, App
c 712	13.8	72.6	601	3	US-09-949-016-149206	Sequence 149206,	c 785	13.8	72.6	1424	3	US-09-590-751A-464	Sequence 464, App
c 713	13.8	72.6	601	3	US-09-949-016-150488	Sequence 150488,	c 786	13.8	72.6	1424	3	US-09-551-621-464	Sequence 464, App
c 714	13.8	72.6	601	3	US-09-949-016-150489	Sequence 150489,	c 787	13.8	72.6	1424	3	US-10-076-622-464	Sequence 464, App
c 715	13.8	72.6	601	3	US-09-949-016-150503	Sequence 150503,	c 788	13.8	72.6	1424	4	US-10-124-805-464	Sequence 464, App
c 716	13.8	72.6	601	3	US-09-949-016-152812	Sequence 152812,	c 789	13.8	72.6	1424	3	US-09-663-600A-49	Sequence 49, Appl
c 717	13.8	72.6	601	3	US-09-949-016-160355	Sequence 160355,	c 790	13.8	72.6	1432	3	US-08-887-534A-50	Sequence 50, Appl
c 718	13.8	72.6	601	3	US-09-949-016-164435	Sequence 164435,	c 791	13.8	72.6	1452	3	US-09-527-431-50	Sequence 50, Appl
c 719	13.8	72.6	601	3	US-09-949-016-164436	Sequence 164436,	c 792	13.8	72.6	1452	3	US-09-446-861-50	Sequence 50, Appl
c 720	13.8	72.6	601	3	US-09-949-016-171192	Sequence 171192,	c 793	13.8	72.6	1452	3	US-09-107-532A-2530	Sequence 2530, Ap
c 721	13.8	72.6	601	3	US-09-949-016-171322	Sequence 171322,	c 794	13.8	72.6	1482	3	US-09-166-350-31	Sequence 31, Appl
c 722	13.8	72.6	601	3	US-09-949-016-178132	Sequence 178132,	c 795	13.8	72.6	1536	3	US-09-489-039A-3610	Sequence 3610, Ap
c 723	13.8	72.6	601	3	US-09-949-016-178452	Sequence 178452,	c 796	13.8	72.6	1542	3	US-09-248-796A-11563	Sequence 11563, A
c 724	13.8	72.6	601	3	US-09-949-016-197046	Sequence 197046,	c 797	13.8	72.6	1578	3	US-09-543-681A-1714	Sequence 1714, Ap
c 725	13.8	72.6	601	3	US-09-949-016-197666	Sequence 197666,	c 798	13.8	72.6	1587	3	US-09-830-230A-587	Sequence 587, App
c 726	13.8	72.6	601	3	US-09-949-016-197667	Sequence 197667,	c 799	13.8	72.6	1599	3	US-09-107-532A-2530	Sequence 2530, Ap
c 727	13.8	72.6	601	3	US-09-949-016-199594	Sequence 199594,	c 800	13.8	72.6	1661	3	US-09-166-350-31	Sequence 31, Appl
c 728	13.8	72.6	601	3	US-09-949-016-200068	Sequence 200068,	c 801	13.8	72.6	1661	3	US-09-220-132-36	Sequence 36, Appl
c 729	13.8	72.6	601	3	US-09-949-016-200068	Sequence 200068,	c 802	13.8	72.6	1661	3	US-09-023-655-1340	Sequence 1340, Ap
c 730	13.8	72.6	601	3	US-09-949-016-205133	Sequence 205133,	c 803	13.8	72.6	1729	3	US-09-620-405B-466	Sequence 466, App
c 731	13.8	72.6	601	3	US-09-949-002-4265	Sequence 4265, Ap	c 804	13.8	72.6	1729	3	US-09-433-826B-466	Sequence 466, App
c 732	13.8	72.6	601	3	US-09-949-002-4266	Sequence 4266, Ap	c 805	13.8	72.6	1729	3	US-09-604-287A-466	Sequence 466, App
c 733	13.8	72.6	601	3	US-09-949-002-5165	Sequence 5165, Ap	c 806	13.8	72.6	1729	3	US-09-834-759-466	Sequence 466, App
c 734	13.8	72.6	601	3	US-09-949-002-9165	Sequence 9165, Ap	c 807	13.8	72.6	1729	3	US-09-590-751A-466	Sequence 466, App
c 735	13.8	72.6	639	3	US-08-956-171E-761	Sequence 761, App	c 808	13.8	72.6	1729	3	US-09-551-621-466	Sequence 466, App
c 736	13.8	72.6	639	3	US-08-781-986A-761	Sequence 761, App	c 809	13.8	72.6	1729	3	US-09-551-621A-466	Sequence 466, App
c 737	13.8	72.6	687	3	US-09-710-279-1861	Sequence 1861, Ap	c 810	13.8	72.6	1729	3	US-10-076-622-466	Sequence 466, App
c 738	13.8	72.6	720	3	US-09-107-532A-1520	Sequence 1520, Ap	c 811	13.8	72.6	1729	4	US-10-124-805-466	Sequence 466, App
c 739	13.8	72.6	722	3	US-09-621-976-17256	Sequence 17256, A	c 812	13.8	72.6	1750	3	US-08-949-770-1	Sequence 1, Appl
c 740	13.8	72.6	732	3	US-09-134-001C-1523	Sequence 1523, Ap	c 813	13.8	72.6	1779	3	US-09-489-039A-3610	Sequence 3610, Ap
c 741	13.8	72.6	751	3	US-08-956-171E-892	Sequence 892, App	c 814	13.8	72.6	1792	3	US-09-270-767-15162	Sequence 2, Appl
c 742	13.8	72.6	751	3	US-08-781-986A-892	Sequence 892, App	c 815	13.8	72.6	1813	3	US-09-558-338-2	Sequence 2, Appl
c 743	13.8	72.6	801	3	US-09-134-000C-3287	Sequence 3287, Ap	c 816	13.8	72.6	1861	3	US-09-743-237-2	Sequence 2, Appl
c 744	13.8	72.6	810	3	US-09-583-110-328	Sequence 328, App	c 817	13.8	72.6	1861	3	US-10-876-841A-2	Sequence 2, Appl
c 745	13.8	72.6	810	3	US-09-769-787-293	Sequence 293, App	c 818	13.8	72.6	1884	3	US-09-593-110-2476	Sequence 203, App
c 746	13.8	72.6	811	3	US-09-270-767-8314	Sequence 8314, Ap	c 819	13.8	72.6	1884	3	US-09-769-787-203	Sequence 2476, Ap
c 747	13.8	72.6	811	3	US-09-270-767-2356	Sequence 2356, A	c 820	13.8	72.6	1907	3	US-08-924-629C-46	Sequence 46, Appl
c 748	13.8	72.6	840	3	US-09-107-433-122	Sequence 122, App	c 821	13.8	72.6	1914	3	US-09-107-433-566	Sequence 566, App
c 749	13.8	72.6	885	5	US-10-770-127-193	Sequence 193, App	c 822	13.8	72.6	1914	3	US-09-107-433-1788	Sequence 1788, Ap
c 750	13.8	72.6	887	3	US-09-270-767-14919	Sequence 14919, A	c 823	13.8	72.6	1946	3	US-09-042-225-5	Sequence 5, Appl
c 751	13.8	72.6	900	5	US-10-770-127-184	Sequence 184, App	c 824	13.8	72.6	1953	3	US-10-076-622-549	Sequence 549, App
c 752	13.8	72.6	903	3	US-05-393-634-63	Sequence 63, Appl	c 825	13.8	72.6	1953	4	US-10-124-805-549	Sequence 549, App
c 753	13.8	72.6	903	5	US-10-770-127-29	Sequence 29, Appl	c 826	13.8	72.6	1970	4	US-10-094-749-1548	Sequence 1548, Ap

973 13.8 72.6 12383 3 US-09-949-016-13323 Sequence 13323, A
974 13.8 72.6 12383 3 US-09-949-016-13324 Sequence 13324, A
C 975 13.8 72.6 12785 3 US-09-553-690-1 Sequence 1, Appli
C 976 13.8 72.6 13598 3 US-09-949-016-11856 Sequence 11856, A
C 977 13.8 72.6 13599 3 US-09-949-016-15876 Sequence 15876, A
C 978 13.8 72.6 13948 3 US-09-949-016-13023 Sequence 13023, A
C 979 13.8 72.6 14042 3 US-08-652-877-85 Sequence 85, Appl
C 980 13.8 72.6 14044 3 US-08-652-877-89 Sequence 89, Appl
C 981 13.8 72.6 14080 3 US-08-652-877-87 Sequence 87, Appl
C 982 13.8 72.6 14083 3 US-08-476-515A-83 Sequence 83, Appl
C 983 13.8 72.6 14086 3 US-08-652-877-83 Sequence 83, Appl
984 13.8 72.6 16134 3 US-09-949-016-13056 Sequence 13056, A
985 13.8 72.6 16935 3 US-09-949-016-15534 Sequence 15534, A
C 986 13.8 72.6 17069 5 US-09-543-679A-2752 Sequence 2752, Ap
C 987 13.8 72.6 17132 3 US-09-949-016-15361 Sequence 15361, A
C 988 13.8 72.6 19272 3 US-09-949-016-13462 Sequence 13462, A
989 13.8 72.6 20099 3 US-09-949-016-13074 Sequence 13074, A
990 13.8 72.6 20661 3 US-09-949-016-13326 Sequence 13326, A
991 13.8 72.6 20661 3 US-09-949-016-13327 Sequence 13327, A
992 13.8 72.6 20661 3 US-09-949-016-13328 Sequence 13328, A
993 13.8 72.6 20661 3 US-09-949-016-13329 Sequence 13329, A
994 13.8 72.6 20661 3 US-09-949-016-13330 Sequence 13330, A
995 13.8 72.6 20661 3 US-09-949-016-13331 Sequence 13331, A
996 13.8 72.6 20662 3 US-09-949-016-15156 Sequence 15156, A
997 13.8 72.6 20681 3 US-09-949-016-14596 Sequence 14596, A
998 13.8 72.6 20681 3 US-09-949-016-14597 Sequence 14597, A
999 13.8 72.6 20681 3 US-09-949-016-14598 Sequence 14598, A
1000 13.8 72.6 20681 3 US-09-949-016-14599 Sequence 14599, A

ALIGNMENTS

RESULT 1
US-09-949-016-29173/c
; Sequence 29173, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29173
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29173

Query Match 100.0%; Score 19; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTG 19
|||||
Db 55 CTTTACTTCATAGCTCTTG 37

RESULT 2
US-09-949-016-126957/c
; Sequence 126957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126957
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126957

Query Match 100.0%; Score 19; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTG 19
|||||
Db 55 CTTTACTTCATAGCTCTTG 37

RESULT 3
US-09-188-930-249/c
; Sequence 249, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-188-930-249

Query Match 100.0%; Score 19; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGCTCTTG 19
|||||
Db 1031 CTTTACTTCATAGCTCTTG 1013

RESULT 4
US-09-312-283C-249/c
; Sequence 249, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use

```
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-249

Query Match      100.0%; Score 19; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
   |||||
Db 1031 CTTTACTTCATAGTCTTTG 1013

RESULT 5
US-09-495-050A-223/c
; Sequence 223, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guesler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2681738CT1
US-09-495-050A-223

Query Match      100.0%; Score 19; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
   |||||
Db 582 CTTTACTTCATAGTCTTTG 564

RESULT 6
US-09-949-016-3613/c
; Sequence 3613, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
US-09-949-016-3613/c

; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3613
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3613

Query Match      100.0%; Score 19; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
   |||||
Db 1158 CTTTACTTCATAGTCTTTG 1140

RESULT 7
US-09-949-016-572/c
; Sequence 572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-572

Query Match      100.0%; Score 19; DB 3; Length 2456;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
   |||||
Db 1312 CTTTACTTCATAGTCTTTG 1294

RESULT 8
US-09-949-016-15355/c
; Sequence 15355, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15355
; LENGTH: 26086
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15355
```

Query Match 100.0%; Score 19; DB 3; Length 26086;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 22954 CTTTACTTCATAGTCTTTG 22936

RESULT 9
US-09-949-016-12314/c
; Sequence 12314, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12314
; LENGTH: 26238
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12314

Query Match 100.0%; Score 19; DB 3; Length 26238;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 23108 CTTTACTTCATAGTCTTTG 23090

RESULT 10
US-09-949-016-173650/c
; Sequence 173650, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173650
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173650

Query Match 91.6%; Score 17.4; DB 3; Length 601;
Best Local Similarity 94.7%; Pred. No. 58;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 221 CTTTACTTCATAGTCTTTG 203

RESULT 11
US-09-949-016-16659
; Sequence 16659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16659
; LENGTH: 50950
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16659

Query Match 91.6%; Score 17.4; DB 3; Length 50950;
Best Local Similarity 94.7%; Pred. No. 89;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 21387 CTTTACTTCATAGTCTTTG 21405

RESULT 12
US-09-513-999C-32761/c
; Sequence 32761, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 32761
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 384
; OTHER INFORMATION: r=a or g
US-09-513-999C-32761

Query Match 86.3%; Score 16.4; DB 3; Length 407;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTTACTTCATAGTCTTTG 19
|||||
Db 40 TTTTACTTCATAGTCTTTG 23

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16745
; LENGTH: 22615
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16745

Query Match      86.3%; Score 16.4; DB 3; Length 22615;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      18145 TTTACTTCATAGTCTTTG 18128

RESULT 16
US-09-949-016-17192
; Sequence 17192, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17192
; LENGTH: 107937
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc.feature
; LOCATION: (1)...(107937)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-17192

Query Match      86.3%; Score 16.4; DB 3; Length 107937;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTT 18
Db      42230 CTTTACTTCATAGTCTTT 42247

RESULT 17
US-09-949-016-17030
; Sequence 17030, Application US/09949016
; Patent No. 6812339

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16805
; LENGTH: 15131
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16805

Query Match      86.3%; Score 16.4; DB 3; Length 15131;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      481 TTTACTTCATAGTCTTTG 464

RESULT 14
US-09-949-016-12148/c
; Sequence 12148, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12148
; LENGTH: 22614
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12148

Query Match      86.3%; Score 16.4; DB 3; Length 22614;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      18145 TTTACTTCATAGTCTTTG 18128

RESULT 15
US-09-949-016-16745/c
; Sequence 16745, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16805
; LENGTH: 15131
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16805

Query Match      86.3%; Score 16.4; DB 3; Length 15131;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      481 TTTACTTCATAGTCTTTG 464
```



```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17030
; LENGTH: 139049
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(139049)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17030
```

```
Query Match      86.3%; Score 16.4; DB 3; Length 139049;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 TTTACTTCATAGTCTTG 19
      |||||
Db      61326 TTTACTTCATGCTTTG 61343
```

```
RESULT 18
US-09-949-016-2159/c
; Sequence 2159, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2159
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2159
```

```
Query Match      84.2%; Score 16; DB 3; Length 2363;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 TTTACTTCATAGTCTT 17
      |||||
Db      780 TTTACTTCATGCTT 765
```

```
RESULT 19
US-09-949-016-13901/c
; Sequence 13901, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13901
; LENGTH: 22206
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(22206)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13901
```

```
Query Match      84.2%; Score 16; DB 3; Length 22206;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 TTTACTTCATAGTCTT 17
      |||||
Db      9697 TTTACTTCATAGTCTT 9692
```

```
RESULT 20
US-09-949-016-12273
; Sequence 12273, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12273
; LENGTH: 130563
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12273
```

```
Query Match      84.2%; Score 16; DB 3; Length 130563;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 TTACTTCATAGTCTTT 18
      |||||
Db      111602 TTACTTCATAGTCTTT 111617
```

```
RESULT 21
US-09-949-016-16050
; Sequence 16050, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16050
; LENGTH: 131379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16050

Query Match 84.2%; Score 16; DB 3; Length 131379;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTT 18
Db 111602 TTACTTCATAGTCTTT 111617

RESULT 22
US-09-949-016-48463
; Sequence 48463, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48463
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48463

Query Match 83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 47 CTTTATTTTCATATTCCTTTG 65

RESULT 23
US-09-949-016-48464
; Sequence 48464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48464
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48464

Query Match 83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 219 CTTTATTTTCATATTCCTTTG 237

RESULT 24
US-09-949-016-48465
; Sequence 48465, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48465
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48465

Query Match 83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
Db 272 CTTTATTTTCATATTCCTTTG 290

RESULT 25
US-09-949-016-48466
; Sequence 48466, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48466
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48466

Query Match      83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 379 CTTTATTTCATATCTTTG 397

RESULT 26
US-09-949-016-48467
; Sequence 48467, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48467
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48467

Query Match      83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 500 CTTTATTTCATATCTTTG 518

RESULT 27
US-09-949-016-48468
; Sequence 48468, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48468
; LENGTH: 601
```

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48468

Query Match      83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 562 CTTTATTTCATATCTTTG 580

RESULT 28
US-09-949-016-121554/c
; Sequence 121554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121554
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-121554

Query Match      83.2%; Score 15.8; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 253 CTTTACTTCATAGTCTTTG 235

RESULT 29
US-09-270-767-12078
; Sequence 12078, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12078
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12078

Query Match      83.2%; Score 15.8; DB 3; Length 1752;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
Db 379 CTCTACTTCATGGTCTTTG 397
```

RESULT 30
US-09-620-312D-1023/c
; Sequence 1023, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radjoe T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pf-Fl_genes Version 1.0
; SEQ ID NO 1023
; LENGTH: 2412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (158)..(544)
US-09-620-312D-1023

Query Match 83.2%; Score 15.8; DB 3; Length 2412;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| ||||| ||||| |||||
Db 746 CTTTATTTCATAGCCTTG 728

RESULT 31
US-09-270-767-15279
; Sequence 15279, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15279
; LENGTH: 2833
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15279

Query Match 83.2%; Score 15.8; DB 3; Length 2833;
Best Local Similarity 89.5%; Pred. No. 3.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| ||||| ||||| |||||
Db 697 CTTTACTTCATGGTCTTG 715

RESULT 32
US-09-949-016-15133/c
; Sequence 15133, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15133
; LENGTH: 27794
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15133

Query Match 83.2%; Score 15.8; DB 3; Length 27794;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||||| ||||| ||||| ||||| |||||
Db 22981 CTTTACTTCTAAGTCTTTG 22963

RESULT 33
US-09-949-016-15990/c
; Sequence 15990, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15990
; LENGTH: 48039
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(48039)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15990

Query Match 83.2%; Score 15.8; DB 3; Length 48039;
Best Local Similarity 89.5%; Pred. No. 4.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
||| ||||| ||||| |||||
Db 36983 CTTGACTTCATATCTTTG 36965

RESULT 34
US-09-949-016-13175/c
; Sequence 13175, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13175
; LENGTH: 58768
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(58768)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13175

Query Match 83.2%; Score 15.8; DB 3; Length 58768;
Best Local Similarity 89.5%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
||| ||||| ||||| |||||
Db 19852 CTTTATTTCATATCTTTG 19834

RESULT 35
US-09-949-016-12009
; Sequence 12009, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12009
; LENGTH: 76221
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12009

Query Match 83.2%; Score 15.8; DB 3; Length 76221;
Best Local Similarity 89.5%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19

Db 10151 CTTAACTTCATAGTCATTG 10169
||| ||||| ||||| |||||

RESULT 36
US-09-949-016-17053
; Sequence 17053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17053
; LENGTH: 76221
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17053

Query Match 83.2%; Score 15.8; DB 3; Length 76221;
Best Local Similarity 89.5%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCATTG 19
||| ||||| ||||| |||||
Db 10151 CTTAACTTCATAGTCATTG 10169

RESULT 37
US-09-949-016-16435/c
; Sequence 16435, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16435
; LENGTH: 258775
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16435

Query Match 83.2%; Score 15.8; DB 3; Length 258775;
Best Local Similarity 89.5%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCATTG 19
||| ||||| ||||| |||||
Db 65440 CTTTATTACATAGTCATTG 65422

RESULT 38

```
US-09-248-796A-154
; Sequence 154, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 154
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-154

Query Match      81.1%; Score 15.4; DB 3; Length 279;
Best Local Similarity 94.1%; Pred. No. 4.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
Db      156 TTTAGTTCATAGTCTTT 172

RESULT 39
US-09-397-787-141
; Sequence 141, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 141
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(451)
; OTHER INFORMATION: n = A,T,C or G
US-09-397-787-141

Query Match      81.1%; Score 15.4; DB 3; Length 451;
Best Local Similarity 94.1%; Pred. No. 4.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      3 TTACTTCATAGTCTTTG 19
Db      404 TTACTTCAAGTCTTTG 420

RESULT 40
US-09-949-016-21055/c
; Sequence 21055, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21055
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21055

Query Match      81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTT 17
Db      407 CTTTACTTCATAGTCTT 391

RESULT 41
US-09-949-016-21056/c
; Sequence 21056, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21056
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21056

Query Match      81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTT 17
Db      407 CTTTACTTCATAGTCTT 391

RESULT 42
US-09-949-016-21057/c
; Sequence 21057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21057
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21057

Query Match      81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTT 17
Db 288 CTTTACTTCATAGTCCTT 272

RESULT 43
US-09-949-016-21058/c
; Sequence 21058, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21058
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21058

Query Match      81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTT 17
Db 79 CTTTACTTCATAGTCCTT 63

RESULT 44
US-09-949-016-60155/c
; Sequence 60155, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60155
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-60155

Query Match      81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTTACTTCATAGTCCTT 18
Db 423 TTTTACTTCATAGTCCTT 407

RESULT 45
US-09-949-016-77977/c
; Sequence 77977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77977
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77977

Query Match      81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTT 17
Db 412 CTTTACTTCATAGTCCTT 396

RESULT 46
US-09-949-016-77978/c
; Sequence 77978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77978
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77978
```


US-09-949-016-77978

Query Match 81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTT 17
|||||
Db 407 CTTTACTTCATAGTCTT 391

RESULT 47

US-09-949-016-77979/c
; Sequence 77979, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77979
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77979

Query Match 81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTT 17
|||||
Db 288 CTTTACTTCATAGTCTT 272

RESULT 48

US-09-949-016-77980/c
; Sequence 77980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77980
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77980

Query Match 81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTT 17
|||||
Db 79 CTTTACTTCATAGTCTT 63

RESULT 49

US-09-949-016-144809/c
; Sequence 144809, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144809
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-144809

Query Match 81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTTACTTCATAGTCTTT 18
|||||
Db 162 TTTTACTTCATAGTCTT 146

RESULT 50

US-09-949-016-162942/c
; Sequence 162942, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 162942
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-162942

Query Match 81.1%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
|||||
Db 508 TTACTTCATAGTCTTTG 492

```

RESULT 51
US-09-949-002-8840
; Sequence 8840, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 8840
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-8840

Query Match      81.1%;   Score 15.4;   DB 3;   Length 601;
Best Local Similarity 94.1%;   Pred. No. 5.1e+02;
Matches 16;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

Qy  3  TTACTTCATAGTCTTTG 19
      | | | | | | | | | |
Db  268  TCACTTCATAGTCTTTG 284

RESULT 52
US-09-614-221A-382/c
; Sequence 382, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 382
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-382

Query Match      81.1%;   Score 15.4;   DB 3;   Length 2163;
Best Local Similarity 94.1%;   Pred. No. 5.7e+02;
Matches 16;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

Qy  3  TTACTTCATAGTCTTTG 19
      | | | | | | | | | |
Db  1662  TTTCTTCATAGTCTTTG 1646

RESULT 53
US-10-104-047-659/c
; Sequence 659, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25

```

Qy 2 TTTACTTCATAGTCCTT 18
|||||
Db 2446 TTTACTTCATAGCCTTT 2430

RESULT 55

US-08-781-986A-295/c
; Sequence 295, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: P8248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 295:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2817 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-295

Query Match 81.1%; Score 15.4; DB 3; Length 2817;
Best Local Similarity 94.1%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCCTT 18
|||||
Db 2446 TTTACTTCATAGCCTTT 2430

RESULT 56

US-09-248-528-5
; Sequence 5, Application US/09248528C
; Patent No. 6153415
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-401
; CURRENT APPLICATION NUMBER: US/09/248,528C
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 60/083,485
; EARLIER FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: BR449
; FEATURE:
; NAME/KEY: gene
; LOCATION: (432)..(1475)
; OTHER INFORMATION: Amidase gene
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1606)..(2292)
; OTHER INFORMATION: Nitrile hydratase beta subunit
; FEATURE:
; NAME/KEY: gene
; LOCATION: (2321)..(2962)
; OTHER INFORMATION: Nitrile hydratase alpha subunit
; FEATURE:
; NAME/KEY: gene
; LOCATION: (2981)..(3283)
; OTHER INFORMATION: ORF 1
US-09-248-528-5

Query Match 81.1%; Score 15.4; DB 3; Length 3299;
Best Local Similarity 94.1%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTT 17
|||||
Db 193 CTTTACTTCCTTAGTCCTT 209

RESULT 57

US-09-549-108-5
; Sequence 5, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
US-09-549-108-5

; SEQ ID NO 5
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: BR449
; FEATURE:
; NAME/KEY: gene
; LOCATION: (432)..(1475)
; OTHER INFORMATION: Amidase gene
; NAME/KEY: gene
; LOCATION: (1606)..(2292)
; OTHER INFORMATION: Nitrile hydratase beta subunit
; NAME/KEY: gene
; LOCATION: (2321)..(2962)
; OTHER INFORMATION: Nitrile hydratase alpha subunit
; NAME/KEY: gene
; LOCATION: (2981)..(3283)
; OTHER INFORMATION: ORF 1
US-09-549-108-5

Query Match 81.1%; Score 15.4; DB 3; Length 3299;
Best Local Similarity 94.1%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 CTTTACTTCATAGTCTT 17
    ||||| |||||
Db 193 CTTTACTTCATAGTCTT 209

RESULT 58
US-09-549-111-5
; Sequence 5, Application US/09549111
; Patent No. 6228633
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-489
; CURRENT APPLICATION NUMBER: US/09/549,111
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: BR449
; FEATURE:
; NAME/KEY: gene
; LOCATION: (432)..(1475)
; OTHER INFORMATION: Amidase gene
; NAME/KEY: gene
; LOCATION: (1606)..(2292)
; OTHER INFORMATION: Nitrile hydratase beta subunit
; NAME/KEY: gene
; LOCATION: (2321)..(2962)
; OTHER INFORMATION: Nitrile hydratase alpha subunit
; NAME/KEY: gene
; LOCATION: (2981)..(3283)
; OTHER INFORMATION: ORF 1
US-09-549-111-5

Query Match 81.1%; Score 15.4; DB 3; Length 3299;
Best Local Similarity 94.1%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTT 17
    ||||| |||||
Db 193 CTTTACTTCATAGTCTT 209

RESULT 60
US-09-550-394-5
; Sequence 5, Application US/09550394
; Patent No. 6287828
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-488
; CURRENT APPLICATION NUMBER: US/09/550,394
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: BR449
; FEATURE:
; NAME/KEY: gene
; LOCATION: (432)..(1475)
; OTHER INFORMATION: Amidase gene
; NAME/KEY: gene
; LOCATION: (1606)..(2292)
; OTHER INFORMATION: Nitrile hydratase beta subunit
; NAME/KEY: gene
; LOCATION: (2321)..(2962)
; OTHER INFORMATION: Nitrile hydratase alpha subunit
; NAME/KEY: gene
; LOCATION: (2981)..(3283)
; OTHER INFORMATION: ORF 1
US-09-550-394-5

Query Match 81.1%; Score 15.4; DB 3; Length 3299;
Best Local Similarity 94.1%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTT 17
    ||||| |||||
Db 193 CTTTACTTCATAGTCTT 209

RESULT 59
US-09-549-106-5
; Sequence 5, Application US/09549106
; Patent No. 6242242
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-487
; CURRENT APPLICATION NUMBER: US/09/549,106
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3299
```

RESULT 61
US-09-354-147C-6
; Sequence 6, Application US/09354147C
; Patent No. 6573067
; GENERAL INFORMATION:
; APPLICANT: Dib-Hajj, Sulayman
; APPLICANT: Waxman, Stephen G.
; TITLE OF INVENTION: Modulation of Sodium Channels in Dorsal Root Ganglia
; FILE REFERENCE: 44574-5004-01-US
; CURRENT APPLICATION NUMBER: US/09/354,147C
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US 60/072,990
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: US 60/109,402
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US99/02008
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3701
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3699)
; OTHER INFORMATION: partial human NaN cDNA sequence
; NAME/KEY: unsure
; LOCATION: (922)
; OTHER INFORMATION: Y = c or t. Xaa at amino acid position 308 is
; OTHER INFORMATION: Leu.
US-09-354-147C-6

Query Match 81.1%; Score 15.4; DB 3; Length 3701;
Best Local Similarity 94.1%; Pred. No. 6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
|||||
Db 3260 TTTACTTCGTAAGTCTTT 3276
|||||

RESULT 62
US-08-927-939A-30/c
; Sequence 30, Application US/08927939A
; Patent No. 6989435
; GENERAL INFORMATION:
; APPLICANT: Grainger, David J.
; APPLICANT: Tatalick, Lauren Marie
; TITLE OF INVENTION: Compounds and methods to inhibit or
; TITLE OF INVENTION: augment an inflammatory response.
; FILE REFERENCE: 1543.001US1
; CURRENT APPLICATION NUMBER: US/08/927,939A
; CURRENT FILING DATE: 1997-11-09
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 4259
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2070)...(2130)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2669)...(2795)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2990)...(3079)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3491)...(3506)
US-08-927-939A-30

Query Match 81.1%; Score 15.4; DB 4; Length 4259;
Best Local Similarity 94.1%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
|||||
Db 1612 TTTACTTCATAGGCTTT 1596
|||||

RESULT 63
US-09-354-147C-41
; Sequence 41, Application US/09354147C
; Patent No. 6573067
; GENERAL INFORMATION:
; APPLICANT: Dib-Hajj, Sulayman
; APPLICANT: Waxman, Stephen G.
; TITLE OF INVENTION: Modulation of Sodium Channels in Dorsal Root Ganglia
; FILE REFERENCE: 44574-5004-01-US
; CURRENT APPLICATION NUMBER: US/09/354,147C
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US 60/072,990
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: US 60/109,402
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US99/02008
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)..(5403)
; OTHER INFORMATION: full length cDNA sequence for human NaN
US-09-354-147C-41

Query Match 81.1%; Score 15.4; DB 3; Length 5860;
Best Local Similarity 94.1%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
|||||
Db 3887 TTTACTTCGTAAGTCTTT 3903
|||||

RESULT 64
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Query Match 81.1%; Score 15.4; DB 3; Length 7679;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
Db 2839 TTACTTCAAGTCTTTG 2855

RESULT 65

US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g31396
; US-09-023-655-1289

Query Match 81.1%; Score 15.4; DB 3; Length 7680;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
Db 2839 TTACTTCAAGTCTTTG 2855

RESULT 66

US-09-543-679A-2698
; Sequence 2698, Application US/09543679A
; Patent No. 7034007
; GENERAL INFORMATION:
; APPLICANT: NYCE, Jonathan W.
; TITLE OF INVENTION: LOW ADENOSINE ANTI-SENSE OLIGONUCLEOTIDE,
; COMPOSITIONS, KIT & METHOD FOR TREATMENT

; OF AIRWAY DISORDERS ASSOCIATED WITH
; BRONCHOCONSTRICTION, LUNG INFLAMMATION,
; NUMBER OF SEQUENCES: 3111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: EPIGENESIS PHARMACEUTICALS, INC.
; STREET: 7 Clarke Drive
; CITY: Cranbury
; STATE: NJ
; COUNTRY: USA
; ZIP: 08512
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/543,679A
; FILING DATE: 13-Apr-2000
; CLASSIFICATION: UNKNOWN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/127,958
; FILING DATE: 1998-08-03
; ATTORNEY/AGENT INFORMATION:
; NAME: Ansel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: EPI-0067191b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-409-3035
; TELEFAX: 413-254-9245
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2698:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2698

US-09-543-679A-2698

Query Match 81.1%; Score 15.4; DB 5; Length 7680;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
Db 2839 TTACTTCAAGTCTTTG 2855

RESULT 67

PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857

; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
PCT-US95-09819-6

Query Match 81.1%; Score 15.4; DB 7; Length 7680;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
Db 2839 TTACTTCAAGTCTTTG 2855

RESULT 68
US-08-259-569-16
; Sequence 16, Application US/08259569
; Patent No. 5679320
; GENERAL INFORMATION:
; APPLICANT: Vogel, Tikva
; APPLICANT: Levanon, Avigdor
; APPLICANT: Werber, Moshe
; APPLICANT: Guy, Rachel
; APPLICANT: Panet, Amos
; APPLICANT: Hartman, Jacob
; APPLICANT: Shaked, Hadassa
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,569
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 703,842
; FILING DATE: 21-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB
; TELEPHONE: 212-664-0525
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7705 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5..7681
US-08-259-569-16

Query Match 81.1%; Score 15.4; DB 2; Length 7705;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
Db 2839 TTACTTCAAGTCTTTG 2855

RESULT 69
US-08-826-885-16
; Sequence 16, Application US/08826885
; Patent No. 5869616
; GENERAL INFORMATION:
; APPLICANT: Vogel, Tikva
; APPLICANT: Levanon, Avigdor
; APPLICANT: Werber, Moshe
; APPLICANT: Guy, Rachel
; APPLICANT: Panet, Amos
; APPLICANT: Hartman, Jacob
; APPLICANT: Shaked, Hadassa
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,885
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/703,842
; FILING DATE: 21-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB
; TELEPHONE: 212-664-0525
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7705 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5..7681

US-08-826-885-16

Query Match 81.1%; Score 15.4; DB 2; Length 7705;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
DB 2839 TTACTTCAAAGTCTTTG 2855

RESULT 70

5455158-2
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO: 2;
; LENGTH: 7705
5455158-2

Query Match 81.1%; Score 15.4; DB 10; Length 7705;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
DB 2839 TTACTTCAAAGTCTTTG 2855

RESULT 71

US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A

; REGISTRATION NUMBER: 33-246
; REFERENCE/DOCKET NUMBER: 92-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7803 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..7346
; US-08-551-356-1

Query Match 81.1%; Score 15.4; DB 2; Length 7803;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
DB 2924 TTACTTCAAAGTCTTTG 2940

RESULT 72

PCT-US93-12687-1
; Sequence 1, Application PC/TUS9312687
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher H.
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12687
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/998,271
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 92-26PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7803 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..7346
; PCT-US93-12687-1

Query Match 81.1%; Score 15.4; DB 7; Length 7803;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19
|||||
Db 2924 TTACTTCAAGTCCTTTG 2940
|||||

RESULT 73
US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Query Match 81.1%; Score 15.4; DB 3; Length 8044;
Best Local Similarity 94.1%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19
|||||
Db 3187 TTACTTCAAGTCCTTTG 3203
|||||

RESULT 74
US-09-543-679A-2699
; Sequence 2699, Application US/09543679A
; Patent No. 7034007
; GENERAL INFORMATION:
; APPLICANT: NYCE, Jonathan W.
; TITLE OF INVENTION: LOW ADENOSINE ANTI-SENSE OLIGONUCLEOTIDE,
; COMPOSITIONS, KIT & METHOD FOR TREATMENT
; OF AIRWAY DISORDERS ASSOCIATED WITH
; BRONCHOCONSTRICTION, LUNG INFLAMMATION,
; NUMBER OF SEQUENCES: 3111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: EPIGENESIS PHARMACEUTICALS, INC.
; STREET: 7 Clarke Drive
; CITY: Cranbury
; STATE: NJ
; COUNTRY: USA
; ZIP: 08512
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/543,679A
; FILING DATE: 13-Apr-2000
; CLASSIFICATION: UNKNOWN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/127,958
; FILING DATE: 1998-08-03
; ATTORNEY/AGENT INFORMATION:
; NAME: Anzel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: EPI-0067191b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-409-3035

; TELEFAX: 413-254-9245
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2699:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14740 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2699
US-09-543-679A-2699

Query Match 81.1%; Score 15.4; DB 5; Length 14740;
Best Local Similarity 94.1%; Pred. No. 6.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19
|||||
Db 9899 TTACTTCAAGTCCTTTG 9915
|||||

RESULT 75
US-09-949-016-16147
; Sequence 16147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16147
; LENGTH: 39937
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16147

Query Match 81.1%; Score 15.4; DB 3; Length 39937;
Best Local Similarity 94.1%; Pred. No. 7.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTT 17
|||||
Db 36637 CTTTACTTCATAGTCCTT 36653
|||||

RESULT 76
US-09-949-016-11904/C
; Sequence 11904, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11904
; LENGTH: 40624
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11904

Query Match 81.1%; Score 15.4; DB 3; Length 40624;
Best Local Similarity 94.1%; Pred. No. 7.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 17
|||||

Db 33515 CTTTACTTCATAGTCTTT 33499

RESULT 77

US-09-949-016-14016/c
; Sequence 14016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14016
; LENGTH: 40625
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14016

Query Match 81.1%; Score 15.4; DB 3; Length 40625;
Best Local Similarity 94.1%; Pred. No. 7.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 17
|||||

Db 33515 CTTTACTTCATAGTCTTT 33499

RESULT 78

US-09-949-016-15836
; Sequence 15836, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15836
; LENGTH: 45587
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15836

Query Match 81.1%; Score 15.4; DB 3; Length 45587;
Best Local Similarity 94.1%; Pred. No. 7.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
|||||

Db 23842 TTTACTTCATAGTCTTT 23858

RESULT 79

US-09-949-016-12674
; Sequence 12674, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12674
; LENGTH: 151256
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)_(151256)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12674

Query Match 81.1%; Score 15.4; DB 3; Length 151256;
Best Local Similarity 94.1%; Pred. No. 8.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TTTACTTCATAGTCTTTG 19
|||||

Db 63211 TTTACTTCATAGTCTTTG 63227

RESULT 80

US-09-949-016-13242
; Sequence 13242, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13242
; LENGTH: 151261
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
US-09-949-016-13242

; LOCATION: (1)...(151261)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-13242

Query Match 81.1%; Score 15.4; DB 3; Length 151261;
Best Local Similarity 94.1%; Pred. NO. 8.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0;

Qy 3 TTACTTCATAGTCTTTG 19
|||
Db 63211 TTCCTTCATAGTCTTTG 63227

RESULT 81

```

US/09-949-002-574/c
; Sequence 574, Application US/09949002
; Patent No. 690016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN
; TITLE OF INVENTION: WITH INFLAMMATO
; TITLE OF INVENTION: AND USES THEREO
; FILE REFERENCE: CU000790
; CURRENT APPLICATION NUMBER: US/09/94
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 1823
; SOFTWARE: FastSeq for Windows Version
; SEQ ID NO 574
; LENGTH: 209631
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(209631)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-574

```

Query Match 81.1%; Score 15.4; DB 3; Length 209631;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0;

QY 3 TTACTTCATAGTCTTG 19
Db 123271 TCACCTTCATAGTCTTG 123255

RESULT 82

```

US-09-949-002-802/c
; Sequence 802, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 802
; LENGTH: 209632
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(209632)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-802

```

Query Match 81.1%; Score 15.4; DB 3; Length 209632;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0;

Qy 3 TTACTTCATAGTCTTTG 19
 | | | | | | | | | |
Db 123271 TCACCTTCATAGTCTTTG 12

RESULT 83

```

US-09-949-016-16480
; Sequence 16480, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16480
; LENGTH: 225127
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/Key: misc feature
; LOCATION: (1)..(225127)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16480

```

```
Query Match      81.1%; Score 15.4; DB 3; Length 225127;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 2 TTTACTTCATAGTCTTT 18
||| ||||| ||||| |||||
Db 3492 TTTTCTTCATAGTCTTT 3508

RESIII.T 84

```

RESULI' 84
US-09-949-016-11868
; Sequence 11868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER-OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 11868
; LENGTH: 300598
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature

```

```
; LOCATION: (1)...(300598)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11868

Query Match      81.1%; Score 15.4; DB 3; Length 300598;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
   |||||
Db 199658 TTTACTTCATAGTATTT 199674

RESULT 85
US-09-949-016-14588
; Sequence 14588, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14588
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14588

Query Match      81.1%; Score 15.4; DB 3; Length 302604;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
   |||||
Db 249664 TTTACTTCATAGTATTT 249680

RESULT 86
US-09-949-016-14589
; Sequence 14589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14589
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14589

Query Match      81.1%; Score 15.4; DB 3; Length 302604;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
   |||||
Db 249664 TTTACTTCATAGTATTT 249680

RESULT 87
US-09-949-016-17119
; Sequence 17119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17119
; LENGTH: 308362
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(308362)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17119

Query Match      81.1%; Score 15.4; DB 3; Length 308362;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
   |||||
Db 249480 TTTACTTCATAGTATTT 249496

RESULT 88
US-09-949-016-13498
; Sequence 13498, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13498
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13498

Query Match      81.1%; Score 15.4; DB 3; Length 302604;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
   |||||
Db 249664 TTTACTTCATAGTATTT 249680
```

```
; SEQ ID NO 13498
; LENGTH: 343352
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(343352)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13498

Query Match      81.1%; Score 15.4; DB 3; Length 343352;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
Db      283626 TTTACTTCATAGTCTTT 283642

RESULT 89
US-09-949-016-14543/c
; Sequence 14543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14543
; LENGTH: 387902
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(387902)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14543

Query Match      81.1%; Score 15.4; DB 3; Length 387902;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
Db      297282 TTTACTTCATAGTCTTT 297266

RESULT 90
US-09-949-016-16297
; Sequence 16297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14543
; LENGTH: 387902
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(387902)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16297

Query Match      81.1%; Score 15.4; DB 3; Length 421183;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
Db      297282 TTTACTTCATAGTCTTT 297266

RESULT 92
US-09-531-120-211
; Sequence 211, Application US/09531120
; Patent No. 6972197
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/09/531,120
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
```

```
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-531-120-211

Query Match      81.1% Score 15.4; DB 4; Length 1082144;
Best Local Similarity 94.1% Pred. No. 7.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTT 18
   |||||
Db  429813 TTTATTTCATAGTCTTT 429829

RESULT 93
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
```



```

; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

```

Query Match      81.1%; Score 15.4; DB 3; Length 1664976;
Best Local Similarity 94.1%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy  2  TTTACTTCATAGTCTTT 18
Db  1303331 TTTACTTCATAGTCTTT 1303315

```

```

RESULT 94
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22

```

```

; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

```

```
/ NAME/KEY: misc_feature
/ LOCATION: (234187)..(234187)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234220)..(234220)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234814)..(234814)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309398)..(309398)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309418)..(309418)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312837)..(312837)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312993)..(312993)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (319226)..(319226)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559167)..(559167)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559241)..(559241)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (600992)..(600992)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (622708)..(622708)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657081)..(657081)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657203)..(657203)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (674435)..(674435)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (682442)..(682442)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (713652)..(713652)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (741684)..(741684)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
```

```
/ LOCATION: (779455)..(779455)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (779576)..(779576)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (855539)..(855539)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (871619)..(871619)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1119881)..(1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1349473)..(1349473)
/ OTHER INFORMATION: n equals a, t, c, or g
```

Query Match 81.1%; Score 15.4; DB 3; Length 1664976;
Best Local Similarity 94.1%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCCTTT 18
|||||
Db 1303331 TTTACTTCATAGTCCTTT 1303315

RESULT 95
US-08-989-299-3/c
; Sequence 3, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: 09/407,427
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-989-299-3

Query Match 78.9%; Score 15; DB 3; Length 2415;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCT 16
|||||
Db 629 TTTACTTCATAGTCT 615

RESULT 96
US-09-407-427-3/c
; Sequence 3, Application US/09407427
; Patent No. 6610497
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THERAPEUTIC
; FILE REFERENCE: MNI-132CP2
; CURRENT APPLICATION NUMBER: US/09/407,427
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/163,648
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-407-427-3

Query Match 78.9%; Score 15; DB 3; Length 2415;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCT 16
|||||
Db 629 TTTACTTCATAGTCT 615

RESULT 97
US-09-635-501-3/c
; Sequence 3, Application US/09635501
; Patent No. 6884771
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L. et al.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THERAPEUTIC
; FILE REFERENCE: MNI-132CP3
; CURRENT APPLICATION NUMBER: US/09/635,501

; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 09/407,427
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/163,648
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-635-501-3

Query Match 78.9%; Score 15; DB 3; Length 2415;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCT 16
|||||
Db 629 TTTACTTCATAGTCT 615

RESULT 98
US-09-163-648-3/c
; Sequence 3, Application US/09163648
; Patent No. 698363
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L. et al.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THERAPEUTIC
; FILE REFERENCE: MNI-132CP
; CURRENT APPLICATION NUMBER: US/09/163,648
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-163-648-3

Query Match 78.9%; Score 15; DB 4; Length 2415;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCT 16
|||||
Db 629 TTTACTTCATAGTCT 615

RESULT 99
US-10-158-847-137/c
; Sequence 137, Application US/10158847
; Patent No. 6592865
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: PFS57
; CURRENT APPLICATION NUMBER: US/10/158,847
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,004
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 137
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:

```
; NAME/KEY: misc feature
; LOCATION: (1707)..(1707)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2702)..(2702)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2749)..(2749)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2757)..(2757)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2788)..(2789)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2819)..(2819)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2835)..(2835)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2856)..(2856)
; OTHER INFORMATION: n equals any amino acid
US-10-158-847-137
```

```
Query Match 78.9%; Score 15; DB 3; Length 2920;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 TTACTTCATAGTCT 16
Db 658 TTACTTCATAGTCT 644
```

RESULT 100

```
US-10-158-825-137/c
; Sequence 137, Application US/10158825
; Patent No. 6900033
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: PF555
; CURRENT APPLICATION NUMBER: US/10/158,825
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/294,976
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 137
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1707)..(1707)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2702)..(2702)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2749)..(2749)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (2757)..(2757)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2788)..(2789)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2819)..(2819)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2835)..(2835)
; OTHER INFORMATION: n equals any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2856)..(2856)
; OTHER INFORMATION: n equals any amino acid
US-10-158-825-137
```

```
Query Match 78.9%; Score 15; DB 3; Length 2920;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 TTACTTCATAGTCT 16
Db 658 TTACTTCATAGTCT 644
```

```
Search completed: June 10, 2006, 15:37:51
Job time : 81.5917 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:26:46 ; Search time 727.7 Seconds
(without alignments) 320.826 Million cell upd

Title: US-10-600-816-32

Perfect score: 19

Sequence: 1 cttacttcatagtctttg 19

Scoring table: IDENTITY NUC

scoring scale: IDENT11_NOC
Gapop 10.0 ; Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Published Applications NA Main:*

1:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US07_PUBCOMB.seq.*
2:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US08_PUBCOMB.seq.*
3:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US09A_PUBCOMB.seq.*
4:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US09B_PUBCOMB.seq.*
5:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US09C_PUBCOMB.seq.*
6:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10A_PUBCOMB.seq.*
7:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10B_PUBCOMB.seq.*
8:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10C_PUBCOMB.seq.*
9:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10D_PUBCOMB.seq.*
10:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10E_PUBCOMB.seq.*
11:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10F_PUBCOMB.seq.*
12:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US10G_PUBCOMB.seq.*
13:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US11A_PUBCOMB.seq.*
14:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US11B_PUBCOMB.seq.*
15:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US11C_PUBCOMB.seq.*
16:	/EMC_Celerra_SIDS3/ptodata/2/pubnpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %		DB	ID	Description
	Score	Match Length			
1	19	100.0	19	8	US-10-600-816-32
2	19	100.0	19	14	US-11-083-784-618880
3	19	100.0	19	15	US-11-101-244-618880
4	19	100.0	42	8	US-10-600-816-30
5	19	100.0	497	6	US-10-066-543-1937
6	19	100.0	552	6	US-10-066-543-1811
7	19	100.0	620	3	US-09-969-034-2322
8	19	100.0	634	3	US-09-969-034-3393
9	19	100.0	642	7	US-10-125-968-701
10	19	100.0	1071	9	US-10-712-615-134
11	19	100.0	1212	3	US-09-866-050A-249
12	19	100.0	1212	6	US-10-152-661-249
13	19	100.0	1228	7	US-10-313-543-233
14	19	100.0	1400	13	US-11-060-756-2418
15	19	100.0	1400	13	US-11-060-756-6690
16	19	100.0	1460	10	US-10-936-626-64
17	19	100.0	1460	10	US-10-938-061-64
					Sequence 32, Appl
					Sequence 618880,
					Sequence 30, Appl
					Sequence 1937, Ap
					Sequence 1811, Ap
					Sequence 2212, Ap
					Sequence 3393, Ap
					Sequence 701, App
					Sequence 134, App
					Sequence 249, App
					Sequence 249, App
					Sequence 223, App
					Sequence 2418, A
					Sequence 6690, A
					Sequence 64, App
					Sequence 64, App

C 91	16.4	86.3	3106	10	US-10-750-185-53364	Sequence 53364, A	C 164	15.8	83.2	613	5	US-09-925-065A-45746	Sequence 45746, A
C 92	16.4	86.3	3106	10	US-10-750-623-53364	Sequence 53364, A	C 165	15.8	83.2	613	12	US-10-301-480-146984	Sequence 146984, A
C 93	16	84.2	19	14	US-11-083-784-470217	Sequence 470217, A	C 166	15.8	83.2	613	12	US-10-301-480-760393	Sequence 760393, A
C 94	16	84.2	19	14	US-11-083-784-470217	Sequence 470217, A	C 167	15.8	83.2	614	4	US-09-925-065A-612190	Sequence 612190, A
C 95	16	84.2	19	15	US-11-101-244-470217	Sequence 470217, A	C 168	15.8	83.2	614	5	US-09-925-065A-612190	Sequence 612190, A
C 96	16	84.2	19	15	US-11-101-244-470217	Sequence 470217, A	C 169	15.8	83.2	619	12	US-10-301-480-235030	Sequence 235030, A
C 97	16	84.2	25	13	US-11-060-756-43358	Sequence 43358, A	C 170	15.8	83.2	619	12	US-10-301-480-838439	Sequence 838439, A
C 98	16	84.2	25	13	US-11-060-756-43361	Sequence 43361, A	C 171	15.8	83.2	620	4	US-09-925-065A-128120	Sequence 128120, A
C 99	16	84.2	50	16	US-11-175-853-39398	Sequence 39398, A	C 172	15.8	83.2	620	4	US-09-925-065A-128120	Sequence 128120, A
C 100	16	84.2	366	12	US-10-301-480-242884	Sequence 242884, A	C 173	15.8	83.2	638	4	US-09-925-065A-901546	Sequence 901546, A
C 101	16	84.2	366	12	US-10-301-480-856293	Sequence 856293, A	C 174	15.8	83.2	638	4	US-09-925-065A-901546	Sequence 901546, A
C 102	16	84.2	367	4	US-09-925-065A-148753	Sequence 148753, A	C 175	15.8	83.2	643	6	US-10-027-632-101884	Sequence 101884, A
C 103	16	84.2	367	5	US-09-925-065A-148753	Sequence 148753, A	C 176	15.8	83.2	643	7	US-10-027-632-101884	Sequence 101884, A
C 104	16	84.2	615	4	US-09-925-065A-789256	Sequence 789256, A	C 177	15.8	83.2	645	4	US-09-925-065A-908942	Sequence 908942, A
C 105	16	84.2	615	5	US-09-925-065A-789256	Sequence 789256, A	C 178	15.8	83.2	645	5	US-09-925-065A-908942	Sequence 908942, A
C 106	16	84.2	624	4	US-09-925-065A-549179	Sequence 549179, A	C 179	15.8	83.2	773	12	US-10-301-480-549761	Sequence 549761, A
C 107	16	84.2	624	5	US-09-925-065A-549179	Sequence 549179, A	C 180	15.8	83.2	773	12	US-10-301-480-549761	Sequence 549761, A
C 108	16	84.2	624	12	US-10-301-480-527129	Sequence 527129, A	C 181	15.8	83.2	773	12	US-10-301-480-1163170	Sequence 1163170, A
C 109	16	84.2	624	12	US-10-301-480-527129	Sequence 527129, A	C 182	15.8	83.2	1236	8	US-10-424-599-58252	Sequence 58252, A
C 110	16	84.2	782	10	US-10-750-185-47810	Sequence 47810, A	C 183	15.8	83.2	1236	9	US-10-487-561-20	Sequence 20, Appl
C 111	16	84.2	782	10	US-10-750-185-47810	Sequence 47810, A	C 184	15.8	83.2	1239	3	US-09-764-864-111	Sequence 111, Appl
C 112	16	84.2	782	10	US-10-750-623-47810	Sequence 47810, A	C 185	15.8	83.2	1259	4	US-09-925-065A-74229	Sequence 74229, A
C 113	16	84.2	1120	5	US-09-925-065A-89841	Sequence 89841, A	C 186	15.8	83.2	1417	5	US-09-925-065A-74229	Sequence 74229, A
C 114	16	84.2	1120	5	US-09-925-065A-89841	Sequence 89841, A	C 187	15.8	83.2	1417	5	US-10-301-480-175468	Sequence 175468, A
C 115	16	84.2	1120	12	US-10-301-480-131082	Sequence 131082, A	C 188	15.8	83.2	1417	12	US-10-301-480-788877	Sequence 788877, A
C 116	16	84.2	1120	12	US-10-301-480-80491	Sequence 80491, A	C 189	15.8	83.2	1599	7	US-10-012-697-1455	Sequence 1455, Ap
C 117	16	84.2	1342	4	US-09-925-065A-89456	Sequence 89456, A	C 190	15.8	83.2	1599	10	US-10-779-543-23465	Sequence 23465, A
C 118	16	84.2	1342	5	US-09-925-065A-89456	Sequence 89456, A	C 191	15.8	83.2	1641	8	US-10-437-963-832	Sequence 832, App
C 119	16	84.2	1342	12	US-10-301-480-190697	Sequence 190697, A	C 192	15.8	83.2	1734	10	US-10-450-763-557	Sequence 557, App
C 120	16	84.2	1342	12	US-10-301-480-804106	Sequence 804106, A	C 193	15.8	83.2	1758	7	US-10-295-027-213	Sequence 213, App
C 121	16	84.2	3737	10	US-10-491-272-88	Sequence 88, Appl	C 194	15.8	83.2	2000	8	US-10-360-238-2386	Sequence 2386, Ap
C 122	16	84.2	103391	10	US-10-981-477-54	Sequence 54, Appl	C 195	15.8	83.2	2000	13	US-11-097-143-28058	Sequence 28058, A
C 123	16	84.2	127567	10	US-10-737-082-47	Sequence 47, Appl	C 196	15.8	83.2	2165	13	US-11-097-143-5477	Sequence 5477, Ap
C 124	15.8	83.2	19	14	US-11-083-784-1167118	Sequence 1167118, A	C 197	15.8	83.2	2175	6	US-10-363-829-132	Sequence 132, App
C 125	15.8	83.2	19	15	US-11-083-784-1167118	Sequence 1167118, A	C 198	15.8	83.2	2412	7	US-10-037-270-1023	Sequence 1023, Ap
C 126	15.8	83.2	440	4	US-09-925-065A-358198	Sequence 358198, A	C 199	15.8	83.2	2412	10	US-10-117-722-1023	Sequence 1023, Ap
C 127	15.8	83.2	440	5	US-09-925-065A-358198	Sequence 358198, A	C 200	15.8	83.2	2412	10	US-10-122-851-1023	Sequence 1023, Ap
C 128	15.8	83.2	442	12	US-10-301-480-429467	Sequence 429467, A	C 201	15.8	83.2	2438	9	US-10-357-930-24567	Sequence 24567, A
C 129	15.8	83.2	442	12	US-10-301-480-1042876	Sequence 1042876, A	C 202	15.8	83.2	2438	9	US-10-357-930-24567	Sequence 24567, A
C 130	15.8	83.2	442	8	US-10-424-599-13969	Sequence 13969, A	C 203	15.8	83.2	2441	13	US-11-097-143-23815	Sequence 23815, A
C 131	15.8	83.2	474	4	US-09-925-065A-781554	Sequence 781554, A	C 204	15.8	83.2	2457	8	US-10-424-599-7103	Sequence 7103, Ap
C 132	15.8	83.2	474	4	US-09-925-065A-842696	Sequence 842696, A	C 205	15.8	83.2	2457	10	US-10-450-763-22320	Sequence 22320, A
C 133	15.8	83.2	474	5	US-09-925-065A-781554	Sequence 781554, A	C 206	15.8	83.2	2935	8	US-10-437-963-48256	Sequence 48256, A
C 134	15.8	83.2	474	5	US-09-925-065A-842696	Sequence 842696, A	C 207	15.8	83.2	2935	8	US-10-437-963-48256	Sequence 48256, A
C 135	15.8	83.2	480	6	US-10-027-632-283411	Sequence 283411, A	C 208	15.8	83.2	3384	3	US-10-437-963-73087	Sequence 73087, A
C 136	15.8	83.2	480	6	US-10-027-632-283411	Sequence 283411, A	C 209	15.8	83.2	10195	3	US-09-764-864-1600	Sequence 1600, Ap
C 137	15.8	83.2	480	7	US-10-027-632-283411	Sequence 283411, A	C 210	15.8	83.2	12462	13	US-11-097-143-5476	Sequence 5476, Ap
C 138	15.8	83.2	480	7	US-10-027-632-283411	Sequence 283411, A	C 211	15.8	83.2	12462	13	US-11-097-143-28057	Sequence 28057, A
C 139	15.8	83.2	500	4	US-09-925-065A-561856	Sequence 561856, A	C 212	15.8	83.2	12462	9	US-10-322-696-133	Sequence 133, App
C 140	15.8	83.2	500	4	US-09-925-065A-561857	Sequence 561857, A	C 213	15.8	83.2	27393	8	US-10-810-788A-6	Sequence 6, Appli
C 141	15.8	83.2	500	5	US-09-925-065A-561857	Sequence 561857, A	C 214	15.8	83.2	87394	9	US-10-987-384-1	Sequence 1, Appli
C 142	15.8	83.2	500	5	US-09-925-065A-561857	Sequence 561857, A	C 215	15.8	83.2	117329	10	US-10-987-384-1	Sequence 9, Appli
C 143	15.8	83.2	573	10	US-10-972-079-76087	Sequence 76087, A	C 216	15.8	83.2	127678	10	US-10-461-862-9	Sequence 9, Appli
C 144	15.8	83.2	588	4	US-09-925-065A-880411	Sequence 880411, A	C 217	15.8	83.2	169739	8	US-10-450-826-93	Sequence 93, Appl
C 145	15.8	83.2	588	5	US-09-925-065A-880411	Sequence 880411, A	C 218	15.8	83.2	207835	15	US-11-121-086-39	Sequence 39, Appl
C 146	15.8	83.2	593	4	US-09-925-065A-632752	Sequence 632752, A	C 219	15.8	83.2	207835	15	US-11-121-086-40	Sequence 40, Appl
C 147	15.8	83.2	593	4	US-09-925-065A-632752	Sequence 632752, A	C 220	15.6	82.1	207835	15	US-11-121-086-40	Sequence 31, Appl
C 148	15.8	83.2	593	5	US-09-925-065A-632752	Sequence 632752, A	C 221	15.8	83.2	209612	8	US-10-322-696-31	Sequence 97, Appl
C 149	15.8	83.2	593	5	US-09-925-065A-632753	Sequence 632753, A	C 222	15.8	83.2	233060	6	US-10-087-192-97	Sequence 97, Appl
C 150	15.8	83.2	593	12	US-10-301-480-335909	Sequence 335909, A	C 223	15.8	83.2	337344	9	US-10-388-838-58	Sequence 58, Appl
C 151	15.8	83.2	593	12	US-10-301-480-949318	Sequence 949318, A	C 224	15.4	81.1	23	9	US-10-600-816-35	Sequence 35, Appl
C 152	15.8	83.2	596	4	US-09-925-065A-896407	Sequence 896407, A	C 225	15.4	81.1	23	9	US-10-600-816-35	Sequence 33, Appl
C 153	15.8	83.2	596	5	US-09-925-065A-896407	Sequence 896407, A	C 226	15.4	81.1	42	9	US-10-719-900-404459	Sequence 404459, A
C 154	15.8	83.2	598	12	US-10-301-480-335907	Sequence 335907, A	C 227	15.4	81.1	25	9	US-10-424-599-134281	Sequence 134281, A
C 155	15.8	83.2	598	12	US-10-301-480-949316	Sequence 949316, A	C 228	15.4	81.1	193	16	US-11-098-686-384	Sequence 384, App
C 156	15.8	83.2	600	10	US-10-972-079-10033	Sequence 10033, A	C 229	15.4	81.1	201	15	US-11-124-368A-11591	Sequence 11591, A
C 157	15.8	83.2	600	10	US-10-972-079-10033	Sequence 10033, A	C 230	15.4	81.1	201	15	US-11-124-368A-11591	Sequence 11591, A
C 158	15.8	83.2	602	4	US-09-925-065A-566918	Sequence 566918, A	C 231	15.4	81.1	339	12	US-10-301-480-236382	Sequence 236382, A
C 159	15.8	83.2	602	4	US-09-925-065A-566918	Sequence 566918, A	C 232	15.4	81.1	339	12	US-10-301-480-236382	Sequence 236382, A
C 160	15.8	83.2	602	4	US-09-925-065A-566918	Sequence 566918, A	C 233	15.4	81.1	350	12	US-10-301-480-849791	Sequence 849791, A
C 161	15.8	83.2	602	5	US-09-925-065A-566918	Sequence 566918, A	C 234	15.4	81.1	351	12	US-10-425-115-157970	Sequence 157970, A
C 162	15.8	83.2	602	5	US-09-925-065A-566919	Sequence 566919, A	C 235	15.4	81.1	351	12	US-10-425-115-157970	Sequence 157970, A
C 163	15.8	83.2	613	4	US-09-925-065A-45746	Sequence 45746, A	C 236	15.4	81.1	354	5	US-09-925-065A-141510	Sequence 141510, A

237	15.4	81.1	459	5	US-09-925-065A-184706, Sequence 184706,	310	15.4	81.1	598	6	US-10-027-632-243582	Sequence 243582,
238	15.4	81.1	461	12	US-10-301-480-274978 Sequence 274978,	311	15.4	81.1	598	7	US-10-027-632-243582	Sequence 243582,
239	15.4	81.1	461	12	US-10-301-480-888387 Sequence 888387,	312	15.4	81.1	599	12	US-10-301-480-33908	Sequence 33908,
240	15.4	81.1	482	3	US-09-918-995-29074 Sequence 29074, A	313	15.4	81.1	599	12	US-10-301-480-949317	Sequence 949317,
241	15.4	81.1	494	5	US-09-925-065A-141505 Sequence 141505,	c 314	15.4	81.1	600	10	US-10-972-079-29176	Sequence 29176, A
242	15.4	81.1	494	5	US-09-925-065A-141505 Sequence 141505,	c 315	15.4	81.1	600	10	US-10-972-079-54266	Sequence 54266, A
243	15.4	81.1	498	12	US-10-301-480-236377 Sequence 236377,	c 316	15.4	81.1	600	10	US-10-972-079-54267	Sequence 54267, A
244	15.4	81.1	498	12	US-10-301-480-849786 Sequence 849786,	c 317	15.4	81.1	600	10	US-10-972-079-54268	Sequence 54268, A
245	15.4	81.1	523	16	US-11-128-061-2442 Sequence 2442, Ap	c 318	15.4	81.1	605	7	US-10-125-968-37	Sequence 27, App1
246	15.4	81.1	523	16	US-11-128-061-6084 Sequence 6084, Ap	c 319	15.4	81.1	608	6	US-10-027-632-193290	Sequence 193290,
247	15.4	81.1	523	16	US-11-128-049-2442 Sequence 2442, Ap	c 320	15.4	81.1	608	6	US-10-027-632-193291	Sequence 193291,
248	15.4	81.1	523	16	US-11-128-049-6084 Sequence 6084, Ap	c 321	15.4	81.1	608	7	US-10-027-632-193290	Sequence 193290,
249	15.4	81.1	525	9	US-10-425-115-43128 Sequence 43128, A	c 322	15.4	81.1	608	7	US-10-027-632-193291	Sequence 193291,
250	15.4	81.1	543	12	US-10-301-480-429464 Sequence 429464,	323	15.4	81.1	611	611	US-10-198-846-13671	Sequence 13671, A
251	15.4	81.1	543	12	US-10-301-480-1042873 Sequence 1042873,	324	15.4	81.1	611	12	US-10-301-480-137419	Sequence 137419, A
252	15.4	81.1	544	4	US-09-925-065A-182736 Sequence 182736,	325	15.4	81.1	611	12	US-10-301-480-74820	Sequence 74820, A
253	15.4	81.1	544	4	US-09-925-065A-182737 Sequence 182737,	326	15.4	81.1	611	12	US-10-301-480-688228	Sequence 688228,
254	15.4	81.1	544	4	US-09-925-065A-182738 Sequence 182738,	327	15.4	81.1	611	12	US-10-301-480-688229	Sequence 688229,
255	15.4	81.1	544	5	US-09-925-065A-182736 Sequence 182736,	c 328	15.4	81.1	613	12	US-10-301-480-424040	Sequence 424040,
256	15.4	81.1	544	5	US-09-925-065A-182737 Sequence 182737,	c 329	15.4	81.1	613	12	US-10-301-480-1037449	Sequence 1037449,
257	15.4	81.1	544	5	US-09-925-065A-182738 Sequence 182738,	c 330	15.4	81.1	616	4	US-09-925-065A-873657	Sequence 873657,
258	15.4	81.1	547	4	US-09-925-065A-231181 Sequence 231181,	c 331	15.4	81.1	616	5	US-09-925-065A-873657	Sequence 873657,
259	15.4	81.1	547	4	US-09-925-065A-231182 Sequence 231182,	332	15.4	81.1	619	4	US-09-925-065A-352227	Sequence 352227,
260	15.4	81.1	547	5	US-09-925-065A-231181 Sequence 231181,	333	15.4	81.1	619	5	US-09-925-065A-352227	Sequence 352227,
261	15.4	81.1	547	5	US-09-925-065A-231182 Sequence 231182,	334	15.4	81.1	619	5	US-10-027-632-234150	Sequence 234150,
262	15.4	81.1	549	4	US-09-925-065A-358194 Sequence 358194,	335	15.4	81.1	624	6	US-10-027-632-234151	Sequence 234151,
263	15.4	81.1	549	5	US-09-925-065A-358194 Sequence 358194,	336	15.4	81.1	624	6	US-10-027-632-234152	Sequence 234152,
264	15.4	81.1	549	12	US-10-301-480-314683 Sequence 314683,	337	15.4	81.1	624	7	US-10-027-632-234152	Sequence 234152,
265	15.4	81.1	549	12	US-10-301-480-314684 Sequence 314684,	338	15.4	81.1	624	7	US-10-027-632-234151	Sequence 234151,
266	15.4	81.1	549	12	US-10-301-480-928092 Sequence 928092,	339	15.4	81.1	624	7	US-10-027-632-234152	Sequence 234152,
267	15.4	81.1	549	12	US-10-301-480-928093 Sequence 928093,	340	15.4	81.1	641	6	US-10-027-632-213247	Sequence 213247,
268	15.4	81.1	550	12	US-10-301-480-273267 Sequence 273267,	341	15.4	81.1	641	7	US-10-027-632-213247	Sequence 213247,
269	15.4	81.1	550	12	US-10-301-480-273268 Sequence 273268,	c 342	15.4	81.1	646	4	US-09-925-065A-530674	Sequence 530674,
270	15.4	81.1	550	12	US-10-301-480-273269 Sequence 273269,	c 343	15.4	81.1	646	5	US-09-925-065A-530674	Sequence 530674,
271	15.4	81.1	550	12	US-10-301-480-886676 Sequence 886676,	344	15.4	81.1	646	12	US-10-301-480-70593	Sequence 70593, A
272	15.4	81.1	550	12	US-10-301-480-886677 Sequence 886677,	345	15.4	81.1	646	12	US-10-301-480-70594	Sequence 70594, A
273	15.4	81.1	550	12	US-10-301-480-886678 Sequence 886678,	346	15.4	81.1	646	12	US-10-301-480-684002	Sequence 684002,
274	15.4	81.1	553	4	US-09-925-065A-141509 Sequence 141509,	347	15.4	81.1	646	12	US-10-301-480-684003	Sequence 684003,
275	15.4	81.1	553	4	US-09-925-065A-331260 Sequence 331260,	348	15.4	81.1	659	4	US-09-925-065A-748225	Sequence 748225,
276	15.4	81.1	553	5	US-09-925-065A-141509 Sequence 141509,	349	15.4	81.1	659	5	US-09-925-065A-748225	Sequence 748225,
277	15.4	81.1	553	5	US-09-925-065A-331260 Sequence 331260,	350	15.4	81.1	659	6	US-10-027-632-227096	Sequence 227096,
278	15.4	81.1	558	12	US-10-301-480-236381 Sequence 236381,	351	15.4	81.1	659	6	US-10-027-632-227097	Sequence 227097,
279	15.4	81.1	558	12	US-10-301-480-849790 Sequence 849790,	352	15.4	81.1	659	7	US-10-027-632-227096	Sequence 227096,
280	15.4	81.1	562	4	US-09-925-065A-339092 Sequence 339092,	353	15.4	81.1	659	7	US-10-027-632-227097	Sequence 227097,
281	15.4	81.1	562	4	US-09-925-065A-339093 Sequence 339093,	c 354	15.4	81.1	674	4	US-09-925-065A-879545	Sequence 879545,
282	15.4	81.1	562	5	US-09-925-065A-339092 Sequence 339092,	c 355	15.4	81.1	674	5	US-09-925-065A-879545	Sequence 879545,
283	15.4	81.1	562	5	US-09-925-065A-339093 Sequence 339093,	c 356	15.4	81.1	751	10	US-10-956-157-3668	Sequence 3668, Ap
284	15.4	81.1	573	4	US-09-925-065A-506143 Sequence 506143,	c 357	15.4	81.1	751	10	US-10-956-157-3668	Sequence 3668, Ap
285	15.4	81.1	573	5	US-09-925-065A-506143 Sequence 506143,	c 358	15.4	81.1	751	10	US-10-956-157-3668	Sequence 3668, Ap
286	15.4	81.1	574	12	US-10-301-480-405063 Sequence 405063,	c 359	15.4	81.1	811	6	US-10-027-632-167490	Sequence 167490,
287	15.4	81.1	574	12	US-10-301-480-1018472 Sequence 1018472,	c 360	15.4	81.1	811	7	US-10-027-632-167490	Sequence 167490,
288	15.4	81.1	576	12	US-10-301-480-412149 Sequence 412149,	361	15.4	81.1	852	8	US-10-198-846-7253	Sequence 7253, Ap
289	15.4	81.1	576	12	US-10-301-480-412150 Sequence 412150,	362	15.4	81.1	852	8	US-10-437-363-43732	Sequence 43732, A
290	15.4	81.1	576	12	US-10-301-480-1025558 Sequence 1025558,	363	15.4	81.1	873	7	US-10-212-677-234	Sequence 234, App
291	15.4	81.1	576	12	US-10-301-480-1025559 Sequence 1025559,	364	15.4	81.1	873	7	US-10-361-811-234	Sequence 234, App
292	15.4	81.1	580	12	US-10-301-480-1025559 Sequence 1025559,	365	15.4	81.1	873	7	US-10-369-186-234	Sequence 234, App
293	15.4	81.1	580	12	US-10-301-480-80515 Sequence 80515, A	366	15.4	81.1	873	16	US-11-250-759-234	Sequence 234, App
294	15.4	81.1	580	12	US-10-301-480-693924 Sequence 693924,	367	15.4	81.1	923	12	US-10-301-480-574262	Sequence 574262,
295	15.4	81.1	583	4	US-09-925-065A-321558 Sequence 321558,	c 368	15.4	81.1	923	12	US-10-301-480-1187671	Sequence 1187671,
296	15.4	81.1	583	4	US-09-925-065A-769126 Sequence 769126,	c 369	15.4	81.1	927	8	US-10-282-122A-12470	Sequence 12470, A
297	15.4	81.1	583	4	US-09-925-065A-769127 Sequence 769127,	370	15.4	81.1	938	16	US-11-079-463-4938	Sequence 4938, Ap
298	15.4	81.1	583	5	US-09-925-065A-321558 Sequence 321558,	371	15.4	81.1	978	12	US-10-301-480-600611	Sequence 600611,
299	15.4	81.1	583	5	US-09-925-065A-769126 Sequence 769126,	372	15.4	81.1	978	12	US-10-301-480-600611	Sequence 600611,
300	15.4	81.1	583	5	US-09-925-065A-769127 Sequence 769127,	c 373	15.4	81.1	1053	16	US-10-301-480-1214020	Sequence 1214020,
301	15.4	81.1	583	12	US-10-301-480-396190 Sequence 396190,	374	15.4	81.1	1070	8	US-11-098-686-8788	Sequence 8788, Ap
302	15.4	81.1	583	12	US-10-301-480-1009599 Sequence 1009599,	375	15.4	81.1	1070	8	US-10-425-114-23979	Sequence 23979, Ap
303	15.4	81.1	591	4	US-09-925-065A-257224 Sequence 257224,	376	15.4	81.1	1190	6	US-10-027-632-116546	Sequence 116546,
304	15.4	81.1	591	4	US-09-925-065A-257225 Sequence 257225,	c 377	15.4	81.1	1190	6	US-10-027-632-116546	Sequence 116546,
305	15.4	81.1	591	4	US-09-925-065A-257226 Sequence 257226,	378	15.4	81.1	1291	10	US-10-750-185-35355	Sequence 35355, A
306	15.4	81.1	591	4	US-09-925-065A-257227 Sequence 257227,	c 379	15.4	81.1	1291	10	US-10-750-185-35355	Sequence 35355, A
307	15.4	81.1	591	5	US-09-925-065A-257228 Sequence 257228,	c 380	15.4	81.1	1335	4	US-10-425-115-48629	Sequence 48629, A
308	15.4	81.1	591	5	US-09-925-065A-257229 Sequence 257229,	c 381	15.4	81.1	1335	4	US-09-925-065A-60635	Sequence 60635, A
309	15.4	81.1	591	5	US-09-925-065A-257227 Sequence 257227,	c 382	15.4	81.1	1335	5	US-09-925-065A-60635	Sequence 60635, A

c 383	15.4	81.1	1335	12	US-10-301-480-161873	Sequence 161873,	456	15.4	81.1	7848	9	US-10-741-600-246	Sequence 246, App
c 384	15.4	81.1	1335	12	US-10-301-480-161874	Sequence 161874,	457	15.4	81.1	7848	10	US-10-995-561-111	Sequence 111, App
c 385	15.4	81.1	1335	12	US-10-301-480-775282	Sequence 775282,	458	15.4	81.1	7867	6	US-10-098-841-6	Sequence 6, Appli
c 386	15.4	81.1	1335	12	US-10-301-480-775283	Sequence 775283,	459	15.4	81.1	7912	16	US-11-193-771-24	Sequence 24, Appl
c 387	15.4	81.1	1467	16	US-11-079-463-65	Sequence 65, Appl	460	15.4	81.1	7912	16	US-11-193-789-24	Sequence 24, Appl
c 388	15.4	81.1	1562	6	US-10-198-846-11719	Sequence 11719, A	461	15.4	81.1	7912	16	US-11-193-806-24	Sequence 24, Appl
c 389	15.4	81.1	1624	8	US-10-437-963-83034	Sequence 83034, A	462	15.4	81.1	7912	16	US-11-193-857-24	Sequence 24, Appl
c 390	15.4	81.1	1806	7	US-10-369-493-33384	Sequence 33384, A	463	15.4	81.1	7912	16	US-11-193-561-24	Sequence 24, Appl
c 391	15.4	81.1	1825	3	US-09-814-353-20464	Sequence 20464, A	464	15.4	81.1	7935	9	US-10-741-601-74	Sequence 74, Appl
c 392	15.4	81.1	1941	8	US-10-282-1228-16053	Sequence 16053, A	465	15.4	81.1	7935	9	US-10-741-600-240	Sequence 240, App
c 393	15.4	81.1	2028	10	US-10-858-412-79	Sequence 79, Appl	466	15.4	81.1	7935	9	US-10-995-561-113	Sequence 113, App
c 394	15.4	81.1	2163	9	US-10-793-639-382	Sequence 382, App	467	15.4	81.1	7959	8	US-10-741-601-81	Sequence 81, Appl
c 395	15.4	81.1	2163	11	US-10-932-182A-78062	Sequence 78062, A	468	15.4	81.1	7959	9	US-10-741-600-249	Sequence 249, App
c 396	15.4	81.1	2197	8	US-10-276-774-839	Sequence 839, App	469	15.4	81.1	7959	10	US-10-995-561-108	Sequence 108, App
c 397	15.4	81.1	2397	7	US-10-104-047-659	Sequence 659, App	470	15.4	81.1	8013	8	US-10-741-601-71	Sequence 71, Appl
c 398	15.4	81.1	2397	16	US-11-072-512-659	Sequence 659, App	471	15.4	81.1	8013	9	US-10-741-600-242	Sequence 242, App
c 399	15.4	81.1	2413	8	US-10-424-599-43603	Sequence 43603, A	472	15.4	81.1	8013	10	US-10-995-561-104	Sequence 104, App
c 400	15.4	81.1	2456	10	US-10-750-185-61032	Sequence 61032, A	473	15.4	81.1	8027	8	US-10-447-161-8	Sequence 8, Appli
c 401	15.4	81.1	2456	10	US-10-750-623-61032	Sequence 61032, A	474	15.4	81.1	8027	8	US-10-734-564-27	Sequence 27, Appl
c 402	15.4	81.1	2494	4	US-09-925-065A-712756	Sequence 712756,	475	15.4	81.1	8027	10	US-10-852-335A-53	Sequence 53, Appl
c 403	15.4	81.1	2494	4	US-09-925-065A-712757	Sequence 712757,	476	15.4	81.1	8027	10	US-10-287-436A-81	Sequence 81, Appl
c 404	15.4	81.1	2494	5	US-09-925-065A-712756	Sequence 712756,	477	15.4	81.1	8044	7	US-10-240-965-121	Sequence 121, App
c 405	15.4	81.1	2494	5	US-09-925-065A-712757	Sequence 712757,	478	15.4	81.1	8044	10	US-10-765-700-135	Sequence 135, App
c 406	15.4	81.1	2532	9	US-10-425-115-106049	Sequence 106049,	479	15.4	81.1	8062	6	US-10-098-841-5	Sequence 5, Appli
c 407	15.4	81.1	2614	4	US-09-925-065A-36691	Sequence 36691, A	480	15.4	81.1	8137	6	US-10-098-841-8	Sequence 8, Appli
c 408	15.4	81.1	2614	5	US-09-925-065A-36691	Sequence 36691, A	481	15.4	81.1	8155	8	US-10-741-601-79	Sequence 79, Appl
c 409	15.4	81.1	2614	12	US-10-301-480-137929	Sequence 137929,	482	15.4	81.1	8155	9	US-10-741-600-247	Sequence 247, App
c 410	15.4	81.1	2614	12	US-10-301-480-751338	Sequence 751338,	483	15.4	81.1	8155	10	US-10-995-561-116	Sequence 116, App
c 411	15.4	81.1	2786	10	US-10-750-185-42848	Sequence 42848, A	484	15.4	81.1	8216	8	US-10-450-763-22270	Sequence 2270, A
c 412	15.4	81.1	2786	10	US-10-750-623-42848	Sequence 42848, A	485	15.4	81.1	8226	8	US-10-741-601-69	Sequence 69, Appl
c 413	15.4	81.1	2817	2	US-08-781-986A-295	Sequence 295, App	486	15.4	81.1	8226	9	US-10-741-600-237	Sequence 237, App
c 414	15.4	81.1	2817	8	US-10-329-624-295	Sequence 295, App	487	15.4	81.1	8226	10	US-10-995-561-107	Sequence 107, App
c 415	15.4	81.1	2873	10	US-10-750-185-24638	Sequence 24638, A	488	15.4	81.1	8230	6	US-10-098-841-7	Sequence 7, Appli
c 416	15.4	81.1	2873	10	US-10-750-623-24638	Sequence 24638, A	489	15.4	81.1	8232	16	US-10-821-234-693	Sequence 693, App
c 417	15.4	81.1	3522	10	US-10-450-763-22266	Sequence 22266, A	490	15.4	81.1	8232	16	US-11-193-771-22	Sequence 22, Appl
c 418	15.4	81.1	3701	7	US-10-388-470-6	Sequence 6, Appli	491	15.4	81.1	8272	16	US-11-193-789-22	Sequence 22, Appl
c 419	15.4	81.1	4013	7	US-10-212-677-267	Sequence 267, App	492	15.4	81.1	8272	16	US-11-193-806-22	Sequence 22, Appl
c 420	15.4	81.1	4013	7	US-10-361-811-267	Sequence 267, App	493	15.4	81.1	8272	16	US-11-193-857-22	Sequence 22, Appl
c 421	15.4	81.1	4013	7	US-10-369-186-267	Sequence 267, App	494	15.4	81.1	8272	16	US-11-193-561-22	Sequence 22, Appl
c 422	15.4	81.1	4013	16	US-11-250-759-267	Sequence 267, App	495	15.4	81.1	8278	8	US-10-741-601-82	Sequence 82, Appl
c 423	15.4	81.1	4259	2	US-08-927-939-30	Sequence 30, Appl	496	15.4	81.1	8278	9	US-10-741-600-250	Sequence 250, App
c 424	15.4	81.1	4259	12	US-10-241-375-30	Sequence 30, Appl	497	15.4	81.1	8278	10	US-10-995-561-106	Sequence 106, App
c 425	15.4	81.1	4385	13	US-11-097-143-8866	Sequence 8866, App	498	15.4	81.1	8332	8	US-10-741-601-73	Sequence 73, Appl
c 426	15.4	81.1	5262	13	US-11-082-454-5	Sequence 5, Appli	499	15.4	81.1	8332	9	US-10-741-600-239	Sequence 239, App
c 427	15.4	81.1	5475	13	US-11-082-454-7	Sequence 7, Appli	500	15.4	81.1	8332	10	US-10-995-561-110	Sequence 110, App
c 428	15.4	81.1	5860	7	US-10-388-470-41	Sequence 41, Appl	501	15.4	81.1	8371	9	US-10-741-601-76	Sequence 76, Appl
c 429	15.4	81.1	6988	8	US-10-236-392-1	Sequence 1, Appli	502	15.4	81.1	8371	9	US-10-741-600-243	Sequence 243, App
c 430	15.4	81.1	7361	8	US-10-236-392-3	Sequence 3, Appli	503	15.4	81.1	8371	10	US-10-995-561-109	Sequence 109, App
c 431	15.4	81.1	7388	16	US-11-193-771-1	Sequence 1, Appli	504	15.4	81.1	8374	16	US-11-193-771-20	Sequence 20, Appl
c 432	15.4	81.1	7388	16	US-11-193-789-1	Sequence 1, Appli	505	15.4	81.1	8374	16	US-11-193-789-20	Sequence 20, Appl
c 433	15.4	81.1	7388	16	US-11-193-806-1	Sequence 1, Appli	506	15.4	81.1	8374	16	US-11-193-806-20	Sequence 20, Appl
c 434	15.4	81.1	7388	16	US-11-193-857-1	Sequence 1, Appli	507	15.4	81.1	8374	16	US-11-193-857-20	Sequence 20, Appl
c 435	15.4	81.1	7388	16	US-11-193-561-1	Sequence 1, Appli	508	15.4	81.1	8374	16	US-11-193-561-20	Sequence 20, Appl
c 436	15.4	81.1	7677	10	US-10-956-157-4995	Sequence 4995, App	509	15.4	81.1	8449	16	US-11-193-771-18	Sequence 18, Appl
c 437	15.4	81.1	7679	10	US-10-831-704-38	Sequence 38, Appl	510	15.4	81.1	8449	16	US-11-193-789-18	Sequence 18, Appl
c 438	15.4	81.1	7680	3	US-09-964-824A-574	Sequence 574, App	511	15.4	81.1	8449	16	US-11-193-806-18	Sequence 18, Appl
c 439	15.4	81.1	7680	6	US-10-171-311-63	Sequence 63, App	512	15.4	81.1	8449	16	US-11-193-857-18	Sequence 18, Appl
c 440	15.4	81.1	7680	7	US-10-236-031B-69	Sequence 69, Appl	513	15.4	81.1	8449	16	US-11-193-561-18	Sequence 18, Appl
c 441	15.4	81.1	7680	7	US-10-374-979-75	Sequence 75, Appl	514	15.4	81.1	8647	16	US-11-193-771-16	Sequence 16, Appl
c 442	15.4	81.1	7680	8	US-10-182-336A-75	Sequence 75, Appl	515	15.4	81.1	8647	16	US-11-193-789-16	Sequence 16, Appl
c 443	15.4	81.1	7680	8	US-10-641-643-1289	Sequence 1289, App	516	15.4	81.1	8647	16	US-11-193-806-16	Sequence 16, Appl
c 444	15.4	81.1	7680	8	US-10-717-597-222	Sequence 222, App	517	15.4	81.1	8647	16	US-11-193-857-16	Sequence 16, Appl
c 445	15.4	81.1	7680	9	US-10-788-792-79	Sequence 79, Appl	518	15.4	81.1	8647	16	US-11-193-561-16	Sequence 16, Appl
c 446	15.4	81.1	7680	9	US-10-477-238A-654	Sequence 654, App	519	15.4	81.1	8815	9	US-10-868-577A-62	Sequence 62, Appl
c 447	15.4	81.1	7680	9	US-10-680-287A-654	Sequence 654, App	520	15.4	81.1	8815	9	US-10-868-549-21	Sequence 21, Appl
c 448	15.4	81.1	7680	9	US-10-278-698-88	Sequence 88, Appl	521	15.4	81.1	8815	16	US-11-193-771-14	Sequence 14, Appl
c 449	15.4	81.1	7680	9	US-10-278-698-603	Sequence 603, Appl	522	15.4	81.1	8815	16	US-11-193-789-14	Sequence 14, Appl
c 450	15.4	81.1	7680	10	US-10-843-641A-5877	Sequence 5877, App	523	15.4	81.1	8815	16	US-11-193-806-14	Sequence 14, Appl
c 451	15.4	81.1	7680	10	US-10-477-173-654	Sequence 654, App	524	15.4	81.1	8815	16	US-11-193-857-14	Sequence 14, Appl
c 452	15.4	81.1	7680	10	US-10-852-335A-52	Sequence 52, Appl	525	15.4	81.1	8815	16	US-11-193-561-14	Sequence 14, Appl
c 453	15.4	81.1	7705	4	US-10-447-161-4	Sequence 4, Appli	526	15.4	81.1	14698	13	US-09-764-868-1429	Sequence 1429, App
c 454	15.4	81.1	7795	6	US-10-084-817-2	Sequence 2, Appli	527	15.4	81.1	17815	13	US-11-082-454-54	Sequence 54, Appl
c 455	15.4	81.1	7848	8	US-10-741-601-78	Sequence 78, Appl	528	15.4	81.1	18031	13	US-11-082-454-57	Sequence 57, Appl

529	15.4	81.1	33794	16	US-11-098-686-8737	Sequence 8737, Ap	c 602	15	78.9	3732	6	US-10-175-746-71	Sequence 71, Appl
530	15.4	81.1	87467	8	US-10-741-601-5634	Sequence 5634, Ap	c 603	15	78.9	3732	6	US-10-176-918-71	Sequence 71, Appl
531	15.4	81.1	87467	9	US-10-741-601-17624	Sequence 17624, A	c 604	15	78.9	3732	6	US-10-176-921-71	Sequence 71, Appl
532	15.4	81.1	87672	10	US-10-995-561-13337	Sequence 13237, A	c 605	15	78.9	3732	6	US-10-137-865-71	Sequence 71, Appl
533	15.4	81.1	90361	9	US-10-719-993-6823	Sequence 6823, Ap	c 606	15	78.9	3732	6	US-10-140-474-71	Sequence 71, Appl
534	15.4	81.1	98606	6	US-10-087-192-31	Sequence 31, Appl	c 607	15	78.9	3732	6	US-10-140-431-71	Sequence 71, Appl
535	15.4	81.1	100000	15	US-11-124-368A-2901	Sequence 2901, Ap	c 608	15	78.9	3732	6	US-10-142-114-71	Sequence 71, Appl
536	15.4	81.1	138821	15	US-11-121-086-80	Sequence 80, Appl	c 609	15	78.9	3732	6	US-10-142-419-71	Sequence 71, Appl
537	15.4	81.1	143947	16	US-11-193-771-37	Sequence 37, Appl	c 610	15	78.9	3732	6	US-10-123-262-71	Sequence 71, Appl
538	15.4	81.1	143947	16	US-11-193-771-37	Sequence 37, Appl	c 611	15	78.9	3732	6	US-10-123-262-71	Sequence 71, Appl
539	15.4	81.1	143947	16	US-11-193-806-37	Sequence 37, Appl	c 612	15	78.9	3732	6	US-10-121-050-71	Sequence 71, Appl
540	15.4	81.1	143947	16	US-11-193-857-37	Sequence 37, Appl	c 613	15	78.9	3732	6	US-10-141-755-71	Sequence 71, Appl
541	15.4	81.1	143947	16	US-11-193-561-37	Sequence 37, Appl	c 614	15	78.9	3732	6	US-10-143-032-71	Sequence 71, Appl
542	15.4	81.1	160170	15	US-11-121-086-32	Sequence 32, Appl	c 615	15	78.9	3732	6	US-10-123-108-71	Sequence 71, Appl
543	15.4	81.1	189539	15	US-11-121-086-16	Sequence 16, Appl	c 616	15	78.9	3732	6	US-10-123-236-71	Sequence 71, Appl
544	15.4	81.1	200000	8	US-10-672-764A-30	Sequence 30, Appl	c 617	15	78.9	3732	6	US-10-123-261-71	Sequence 71, Appl
545	15.4	81.1	200000	10	US-10-984-723-1	Sequence 1, Appl	c 618	15	78.9	3732	6	US-10-140-921-71	Sequence 71, Appl
546	15.4	81.1	352938	8	US-10-322-696-79	Sequence 79, Appl	c 619	15	78.9	3732	6	US-10-140-928-71	Sequence 71, Appl
547	15.4	81.1	1082144	15	US-11-117-187-211	Sequence 211, App	c 620	15	78.9	3732	6	US-10-121-045-71	Sequence 71, Appl
548	15	78.9	19	14	US-11-083-784-470128	Sequence 470128, A	c 621	15	78.9	3732	6	US-10-123-292-71	Sequence 71, Appl
549	15	78.9	19	14	US-11-083-784-470228	Sequence 470228, A	c 622	15	78.9	3732	6	US-10-123-292-71	Sequence 71, Appl
550	15	78.9	19	15	US-11-101-244-470128	Sequence 470128, A	c 623	15	78.9	3732	6	US-10-124-819-71	Sequence 71, Appl
551	15	78.9	19	15	US-11-101-244-470228	Sequence 470228, A	c 624	15	78.9	3732	6	US-10-124-822-71	Sequence 71, Appl
552	15	78.9	25	13	US-11-060-756-43360	Sequence 43360, A	c 625	15	78.9	3732	6	US-10-140-925-71	Sequence 71, Appl
553	15	78.9	158	3	US-09-864-761-27572	Sequence 27572, A	c 626	15	78.9	3732	6	US-10-160-498-71	Sequence 71, Appl
554	15	78.9	279	9	US-10-425-115-88166	Sequence 88166, A	c 627	15	78.9	3732	6	US-10-124-824-71	Sequence 71, Appl
555	15	78.9	476	3	US-09-864-761-10935	Sequence 10935, A	c 628	15	78.9	3732	6	US-10-127-825A-71	Sequence 71, Appl
556	15	78.9	487	3	US-09-770-961-748	Sequence 748, App	c 629	15	78.9	3732	6	US-10-127-829A-71	Sequence 71, Appl
557	15	78.9	591	4	US-09-925-065A-556159	Sequence 556159, A	c 630	15	78.9	3732	6	US-10-127-839A-71	Sequence 71, Appl
558	15	78.9	591	5	US-09-925-065A-556159	Sequence 556159, A	c 631	15	78.9	3732	6	US-10-127-839A-71	Sequence 71, Appl
559	15	78.9	599	7	US-10-029-386-102	Sequence 102, App	c 632	15	78.9	3732	6	US-10-127-901A-71	Sequence 71, Appl
560	15	78.9	623	4	US-09-925-065A-867459	Sequence 867459, A	c 633	15	78.9	3732	6	US-10-128-903A-71	Sequence 71, Appl
561	15	78.9	623	5	US-09-925-065A-867459	Sequence 867459, A	c 634	15	78.9	3732	6	US-10-131-813A-71	Sequence 71, Appl
562	15	78.9	635	12	US-10-301-480-53832	Sequence 53832, A	c 635	15	78.9	3732	6	US-10-131-818A-71	Sequence 71, Appl
563	15	78.9	635	12	US-10-301-480-53833	Sequence 53833, A	c 636	15	78.9	3732	6	US-10-131-823A-71	Sequence 71, Appl
564	15	78.9	635	12	US-10-301-480-667241	Sequence 667241, A	c 637	15	78.9	3732	6	US-10-131-824A-71	Sequence 71, Appl
565	15	78.9	635	12	US-10-301-480-667242	Sequence 667242, A	c 638	15	78.9	3732	6	US-10-131-830A-71	Sequence 71, Appl
566	15	78.9	795	16	US-11-079-463-2905	Sequence 2905, Ap	c 639	15	78.9	3732	6	US-10-131-837A-71	Sequence 71, Appl
567	15	78.9	991	12	US-10-301-480-580216	Sequence 580216, A	c 640	15	78.9	3732	6	US-10-137-872A-71	Sequence 71, Appl
568	15	78.9	991	12	US-10-301-480-1193625	Sequence 1193625, A	c 641	15	78.9	3732	6	US-10-147-500-71	Sequence 71, Appl
569	15	78.9	993	12	US-10-301-480-579953	Sequence 579953, A	c 642	15	78.9	3732	6	US-10-147-502-71	Sequence 71, Appl
570	15	78.9	993	12	US-10-301-480-1186362	Sequence 1186362, A	c 643	15	78.9	3732	6	US-10-147-515-71	Sequence 71, Appl
571	15	78.9	1155	3	US-09-938-842A-1650	Sequence 1650, Ap	c 644	15	78.9	3732	6	US-10-147-517-71	Sequence 71, Appl
572	15	78.9	1155	3	US-09-938-842A-1650	Sequence 1650, Ap	c 645	15	78.9	3732	6	US-10-147-526-71	Sequence 71, Appl
573	15	78.9	1183	6	US-10-138-846-10841	Sequence 10841, A	c 646	15	78.9	3732	6	US-10-147-527-71	Sequence 71, Appl
574	15	78.9	1760	8	US-10-424-599-78779	Sequence 78779, A	c 647	15	78.9	3732	6	US-10-121-041-71	Sequence 71, Appl
575	15	78.9	2911	13	US-11-059-218-3	Sequence 3, Appl	c 648	15	78.9	3732	6	US-10-121-043-71	Sequence 71, Appl
576	15	78.9	2911	3	US-09-969-384-12	Sequence 12, Appl	c 649	15	78.9	3732	6	US-10-121-047-71	Sequence 71, Appl
577	15	78.9	2920	3	US-09-969-384-2	Sequence 2, Appl	c 650	15	78.9	3732	6	US-10-123-215-71	Sequence 71, Appl
578	15	78.9	2920	6	US-10-158-847-137	Sequence 137, App	c 651	15	78.9	3732	6	US-10-123-902-71	Sequence 71, Appl
579	15	78.9	2920	7	US-10-158-825-137	Sequence 137, App	c 652	15	78.9	3732	6	US-10-123-908-71	Sequence 71, Appl
580	15	78.9	2920	8	US-10-158-825-137	Sequence 137, App	c 653	15	78.9	3732	6	US-10-123-909-71	Sequence 71, Appl
581	15	78.9	3324	6	US-10-116-802-172	Sequence 172, App	c 654	15	78.9	3732	6	US-10-123-910-71	Sequence 71, Appl
582	15	78.9	3325	6	US-10-114-893-85	Sequence 85, Appl	c 655	15	78.9	3732	6	US-10-124-813-71	Sequence 71, Appl
583	15	78.9	3334	3	US-09-978-385-1	Sequence 1, Appl	c 656	15	78.9	3732	6	US-10-124-817-71	Sequence 71, Appl
584	15	78.9	3396	3	US-09-999-781-1	Sequence 1, Appl	c 657	15	78.9	3732	6	US-10-125-922-71	Sequence 71, Appl
585	15	78.9	3396	6	US-10-158-847-141	Sequence 141, App	c 658	15	78.9	3732	6	US-10-125-924-71	Sequence 71, Appl
586	15	78.9	3396	7	US-10-158-825-141	Sequence 141, App	c 659	15	78.9	3732	6	US-10-140-860-71	Sequence 71, Appl
587	15	78.9	3396	8	US-10-158-825-141	Sequence 141, App	c 660	15	78.9	3732	6	US-10-142-417-71	Sequence 71, Appl
588	15	78.9	3396	13	US-11-059-218-1	Sequence 1, Appl	c 661	15	78.9	3732	6	US-10-147-519-71	Sequence 71, Appl
589	15	78.9	3405	6	US-10-005-956-569	Sequence 569, App	c 662	15	78.9	3732	6	US-10-157-782-71	Sequence 71, Appl
590	15	78.9	3405	6	US-10-005-956-842	Sequence 842, App	c 663	15	78.9	3732	6	US-10-152-395-71	Sequence 71, Appl
591	15	78.9	3405	10	US-10-356-157-2371	Sequence 2371, Ap	c 664	15	78.9	3732	6	US-10-125-926A-71	Sequence 71, Appl
592	15	78.9	3405	10	US-10-756-149-2496	Sequence 2496, Ap	c 665	15	78.9	3732	6	US-10-125-930A-71	Sequence 71, Appl
593	15	78.9	3405	10	US-10-798-923-4	Sequence 4, Appl	c 666	15	78.9	3732	6	US-10-127-831A-71	Sequence 71, Appl
594	15	78.9	3405	10	US-10-518-599-1	Sequence 1, Appl	c 667	15	78.9	3732	6	US-10-127-837A-71	Sequence 71, Appl
595	15	78.9	3405	16	US-11-091-883-51	Sequence 51, Appl	c 668	15	78.9	3732	6	US-10-127-838B-71	Sequence 71, Appl
596	15	78.9	3474	10	US-10-204-921-37	Sequence 37, Appl	c 669	15	78.9	3732	6	US-10-127-842A-71	Sequence 71, Appl
597	15	78.9	3732	6	US-10-028-072-71	Sequence 71, Appl	c 670	15	78.9	3732	6	US-10-127-843A-71	Sequence 71, Appl
598	15	78.9	3732	6	US-10-140-808-71	Sequence 71, Appl	c 671	15	78.9	3732	6	US-10-127-845A-71	Sequence 71, Appl
599	15	78.9	3732	6	US-10-121-049-71	Sequence 71, Appl	c 672	15	78.9	3732	6	US-10-127-846A-71	Sequence 71, Appl
600	15	78.9	3732	6	US-10-123-904-71	Sequence 71, Appl	c 673	15	78.9	3732	6	US-10-127-848A-71	Sequence 71, Appl
601	15	78.9	3732	6	US-10-140-470-71	Sequence 71, Appl	c 674	15	78.9	3732	6	US-10-127-849A-71	Sequence 71, Appl


```
c 967      15 78.9 3732 7 US-10-137-867-71 Sequence 71, Appl
c 968      15 78.9 3732 7 US-10-140-473-71 Sequence 71, Appl
c 969      15 78.9 3732 7 US-10-140-806-71 Sequence 71, Appl
c 970      15 78.9 3732 7 US-10-140-810-71 Sequence 71, Appl
c 971      15 78.9 3732 7 US-10-140-863-71 Sequence 71, Appl
c 972      15 78.9 3732 7 US-10-141-699-71 Sequence 71, Appl
c 973      15 78.9 3732 7 US-10-141-703-71 Sequence 71, Appl
c 974      15 78.9 3732 7 US-10-141-706-71 Sequence 71, Appl
c 975      15 78.9 3732 7 US-10-141-757-71 Sequence 71, Appl
c 976      15 78.9 3732 7 US-10-141-762-71 Sequence 71, Appl
c 977      15 78.9 3732 7 US-10-142-428-71 Sequence 71, Appl
c 978      15 78.9 3732 7 US-10-142-429-71 Sequence 71, Appl
c 979      15 78.9 3732 7 US-10-142-884-71 Sequence 71, Appl
c 980      15 78.9 3732 7 US-10-143-027-71 Sequence 71, Appl
c 981      15 78.9 3732 7 US-10-143-115-71 Sequence 71, Appl
c 982      15 78.9 3732 7 US-10-144-956-71 Sequence 71, Appl
c 983      15 78.9 3732 7 US-10-144-958-71 Sequence 71, Appl
c 984      15 78.9 3732 7 US-10-145-632-71 Sequence 71, Appl
c 985      15 78.9 3732 7 US-10-145-749-71 Sequence 71, Appl
c 986      15 78.9 3732 7 US-10-145-753-71 Sequence 71, Appl
c 987      15 78.9 3732 7 US-10-145-871-71 Sequence 71, Appl
c 988      15 78.9 3732 7 US-10-145-878-71 Sequence 71, Appl
c 989      15 78.9 3732 7 US-10-146-794-71 Sequence 71, Appl
c 990      15 78.9 3732 7 US-10-147-489-71 Sequence 71, Appl
c 991      15 78.9 3732 7 US-10-147-507-71 Sequence 71, Appl
c 992      15 78.9 3732 7 US-10-147-535-71 Sequence 71, Appl
c 993      15 78.9 3732 7 US-10-147-537-71 Sequence 71, Appl
c 994      15 78.9 3732 7 US-10-152-376-71 Sequence 71, Appl
c 995      15 78.9 3732 7 US-10-152-381-71 Sequence 71, Appl
c 996      15 78.9 3732 7 US-10-152-400-71 Sequence 71, Appl
c 997      15 78.9 3732 7 US-10-153-585-71 Sequence 71, Appl
c 998      15 78.9 3732 7 US-10-157-800-71 Sequence 71, Appl
c 999      15 78.9 3732 7 US-10-157-800-71 Sequence 71, Appl
c1000     15 78.9 3732 7 US-10-157-801-71 Sequence 71, Appl
```

ALIGNMENTS

```
RESULT 1
US-10-600-816-32
; Sequence 32, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RA13, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S.60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-32
```

```
Query Match      100.0%; Score 19; DB 8; Length 19;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 CTTTACTTCATAGCTTTG 19
      |||||
Db      1 CTTTACTTCATAGCTTTG 19
```

RESULT 2

```
US-11-083-784-618880/c
; Sequence 618880, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIOR APPLICATION NUMBER: US/10/714,333
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 618880
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-083-784-618880
```

```
Query Match      100.0%; Score 19; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 CTTTACTTCATAGCTTTG 19
      |||||
Db      19 CTTTACTTCATAGCTTTG 1
```

RESULT 3

```
US-11-101-244-618880/c
; Sequence 618880, Application US/1101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 618880
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-618880
```

```
Query Match      100.0%; Score 19; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 CTTTACTTCATAGCTTTG 19
      |||||
Db      19 CTTTACTTCATAGCTTTG 1
```

RESULT 4
US-10-600-816-30/c
; Sequence 30, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RA13, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR FILING DATE: 2002-06-20
; PRIOR FILING DATE: 2002-06-20
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 30
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-30

Query Match 100.0%; Score 19; DB 8; Length 42;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 42 CTTTACTTCATAGTCTTTG 24

RESULT 5
US-10-066-543-1937/c
; Sequence 1937, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1937
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1937

Query Match 100.0%; Score 19; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 385 CTTTACTTCATAGTCTTTG 367

RESULT 6
US-10-066-543-1811

; Sequence 1811, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1811

Query Match 100.0%; Score 19; DB 6; Length 552;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 113 CTTTACTTCATAGTCTTTG 131

RESULT 7
US-09-969-034-2222
; Sequence 2222, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR FILING DATE: 2001-10-02
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2222
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 491, 500, 517, 530, 545, 547, 553, 558, 570, 599, 620
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-2222

Query Match 100.0%; Score 19; DB 3; Length 620;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19

US-09-866-050A-249

Query Match 100.0%; Score 19; DB 3; Length 1212;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1031 CTTTACTTCATAGTCCTTTG 1013

RESULT 12

US-10-152-661-249/c
; Sequence 249, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152.661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-10-152-661-249

Query Match 100.0%; Score 19; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1031 CTTTACTTCATAGTCCTTTG 1013

RESULT 13

US-10-313-542-223/c
; Sequence 223, Application US/10313542
; Publication No. US20030120057A1
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/10/313.542
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/495,050
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318

; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030120057A1 2681738CT1
US-10-313-542-223

Query Match 100.0%; Score 19; DB 7; Length 1228;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 582 CTTTACTTCATAGTCCTTTG 564

RESULT 14

US-11-060-756-2418/c
; Sequence 2418, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060.756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2418
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-2418

Query Match 100.0%; Score 19; DB 13; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 256 CTTTACTTCATAGTCCTTTG 238

RESULT 15

US-11-060-756-6690/c
; Sequence 6690, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060.756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6690
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-6690

Query Match 100.0%; Score 19; DB 13; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 256 CTTTACTTCATAGTCTTTG 238

RESULT 16

US-10-936-626-64/c
; Sequence 64, Application US/10936626
; Publication No. US20050106644A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-64

Query Match 100.0%; Score 19; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1186 CTTTACTTCATAGTCTTTG 1168
RESULT 17
US-10-938-061-64/c
; Sequence 64, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-64

Query Match 100.0%; Score 19; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1186 CTTTACTTCATAGTCTTTG 1168

RESULT 17

US-10-938-061-64/c
; Sequence 64, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-64

; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie
; APPLICANT: Chuntharapal, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-938-061-64

Query Match 100.0%; Score 19; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1186 CTTTACTTCATAGTCTTTG 1168
RESULT 18
US-10-224-289-5/c
; Sequence 5, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; TITLE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-5

Query Match 100.0%; Score 19; DB 10; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1186 CTTTACTTCATAGTCTTTG 1168

RESULT 18
US-10-224-289-5/c
; Sequence 5, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; TITLE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-5

Query Match 100.0%; Score 19; DB 7; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1178 CTTTACTTCATAGTCTTTG 1160

RESULT 19

US-10-935-190-43/c
; Sequence 43, Application US/10935190
; Publication No. US20050037466A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BURFORD, Neil
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; TITLE OF INVENTION: RECEPTORS AND ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0726 PCT
; CURRENT APPLICATION NUMBER: US/10/935,190
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US/10/031,904
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/145,232; 60/158,578; 60/165,192
; PRIOR FILING DATE: 1999-07-21; 1999-10-07; 1999-11-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 43
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No: 2681738CB1
US-10-935-190-43

Query Match 100.0%; Score 19; DB 9; Length 1619;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1178 CTTTACTTCATAGTCTTTG 1160

RESULT 20

US-10-775-920-13/c
; Sequence 13, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-13

Query Match 100.0%; Score 19; DB 9; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1166 CTTTACTTCATAGTCTTTG 1148

RESULT 21

US-10-224-289-3/c
; Sequence 3, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; FILE REFERENCE: NUCLEIC ACID
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-3

Query Match 100.0%; Score 19; DB 7; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||
Db 1158 CTTTACTTCATAGTCTTTG 1140

RESULT 22

US-10-240-425-405/c
; Sequence 405, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 405
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF095448
US-10-240-425-405

Query Match 100.0%; Score 19; DB 8; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 CTTTACTTCATAGTCTTTG 19
      |||||||
Db      1158 CTTTACTTCATAGTCTTTG 1140

RESULT 23
US-10-775-920-9/c
; Sequence 9, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merigen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-9

Query Match      100.0%; Score 19; DB 9; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCTTTG 19
      |||||||
Db      1158 CTTTACTTCATAGTCTTTG 1140

RESULT 24
US-10-510-507-2/c
; Sequence 2, Application US/10510507
; Publication No. US20050282165A1
; GENERAL INFORMATION:
; APPLICANT: Terrett, Jonathan A
; TITLE OF INVENTION: DIAGNOSIS OF CARCINOMA USING RAIG1 POLYPEPTIDES
; FILE REFERENCE: 2543-1-039PCT/US
; CURRENT APPLICATION NUMBER: US/10/510,507
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: GB0208331.9
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: GB0221538.2
; PRIOR FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-510-507-2

Query Match      100.0%; Score 19; DB 10; Length 2302;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCTTTG 19
      |||||||
Db      1158 CTTTACTTCATAGTCTTTG 1140

RESULT 25
US-10-775-920-12/c
; Sequence 12, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merigen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-12

Query Match      100.0%; Score 19; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCTTTG 19
      |||||||
Db      1162 CTTTACTTCATAGTCTTTG 1144

RESULT 26
US-10-176-847-59/c
; Sequence 59, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-59

Query Match      100.0%; Score 19; DB 6; Length 2316;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCTTTG 19
      |||||||
Db      1164 CTTTACTTCATAGTCTTTG 1146

RESULT 27
US-11-080-991-59/c
; Sequence 59, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-080-991-59

Query Match      100.0%; Score 19; DB 15; Length 2316;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTTG 19
Db      1164 CTTTACTTCATAGTCTTTG 1146

RESULT 28
US-10-775-920-11/c
; Sequence 11, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Merigen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775.920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 2446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-11

Query Match      100.0%; Score 19; DB 9; Length 2446;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTTG 19
Db      1312 CTTTACTTCATAGTCTTTG 1294

RESULT 29
US-10-225-567A-453/c
; Sequence 453, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225.567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 453
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-453

Query Match      100.0%; Score 19; DB 6; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTTG 19
Db      1164 CTTTACTTCATAGTCTTTG 1146
```

```
Db      1312 CTTTACTTCATAGTCTTTG 1294

RESULT 30
US-10-269-909-63/c
; Sequence 63, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIREBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-63

Query Match      100.0%; Score 19; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTTG 19
Db      1312 CTTTACTTCATAGTCTTTG 1294

RESULT 31
US-10-269-909-64/c
; Sequence 64, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIREBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-64

Query Match      100.0%; Score 19; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTTTG 19
Db      1312 CTTTACTTCATAGTCTTTG 1294

RESULT 32
US-10-295-027-619/c
```

```
; Sequence 619, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezl, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 619
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-619

Query Match      100.0%; Score 19; DB 7; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCCTTTG 19
        |||||
Db      1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 33
US-10-600-816-2/c
; Sequence 2, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S.60/407,006
; PRIOR FILING DATE: 2002-08-29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 619
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RAI3 Polymorphic Allele Summary Sequence.
; NAME/KEY: misc feature
; LOCATION: (112)..(112)
; OTHER INFORMATION: wherein "n" equals either G or A.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (364)..(364)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: wherein "n" equals either T or C.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1173)..(1173)
; OTHER INFORMATION: wherein "n" equals either A or G.
; NUMBER OF SEQ ID NOS: 98
US-10-600-816-18

Query Match      100.0%; Score 19; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCCTTTG 19
        |||||
Db      1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 34
US-10-600-816-18/c
; Sequence 18, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S.60/407,006
; PRIOR FILING DATE: 2002-08-29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RAI3 Polymorphic Allele Summary Sequence.
; NAME/KEY: misc feature
; LOCATION: (112)..(112)
; OTHER INFORMATION: wherein "n" equals either G or A.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (364)..(364)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: wherein "n" equals either T or C.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1173)..(1173)
; OTHER INFORMATION: wherein "n" equals either A or G.
; NUMBER OF SEQ ID NOS: 98
US-10-600-816-18
```

Query Match 100.0%; Score 19; DB 8; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 35
US-10-775-920-10/c
; Sequence 10, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-10

Query Match 100.0%; Score 19; DB 9; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 36
US-10-936-626-40/c
; Sequence 40, Application US/10936626
; Publication No. US2005010664A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; CURRENT FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 40
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-40

Query Match 100.0%; Score 19; DB 10; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 37
US-10-938-061-40/c
; Sequence 40, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 40
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-938-061-40

Query Match 100.0%; Score 19; DB 10; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 38
US-11-169-041-32/c
; Sequence 32, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-169-041-32

Query Match 100.0%; Score 19; DB 16; Length 2456;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1312 CTTTACTTCATAGTCCTTTG 1294

RESULT 39
US-10-264-049-834/c
; Sequence 834, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 834
; LENGTH: 2593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-049-834

Query Match 100.0%; Score 19; DB 7; Length 2593;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1424 CTTTACTTCATAGTCCTTTG 1406

RESULT 40
US-10-198-846-10424/c
; Sequence 10424, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10424
; LENGTH: 4239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 4238, 4239
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10424

Query Match 100.0%; Score 19; DB 6; Length 4239;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTTG 19
|||||
Db 1720 CTTTACTTCATAGTCCTTTG 1702

RESULT 41
US-09-925-065A-840554
; Sequence 840554, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 840554
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-840554

Query Match 92.6%; Score 17.6; DB 4; Length 523;
Best Local Similarity 94.4%; Pred. No. 4.1e+02;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTT 18
|||||

Db 403 CTTTACTTMTAGTCTTT 420

RESULT 42

US-09-925-065A-840554
; Sequence 840554, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 840554
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-840554

Query Match 92.6%; Score 17.6; DB 5; Length 523;
Best Local Similarity 94.4%; Pred. No. 4.1e+02;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTACTTMTAGTCTTT 18
|||||:|||||
Db 403 CTTTACTTMTAGTCTTT 420

RESULT 43

US-09-925-065A-916369
; Sequence 916369, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 916369
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-916369

Query Match 91.6%; Score 17.4; DB 4; Length 638;
Best Local Similarity 94.7%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||:|||||
Db 287 CTTTACTTCATAGTCTTTG 305

RESULT 44

US-09-925-065A-916369
; Sequence 916369, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 916369
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-916369

Query Match 91.6%; Score 17.4; DB 5; Length 638;
Best Local Similarity 94.7%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||:|||||
Db 287 CTTTACTTCATAGTCTTTG 305

RESULT 45

US-10-301-480-556032/c
; Sequence 556032, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 556032
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-556032

Query Match 91.6%; Score 17.4; DB 12; Length 997;
Best Local Similarity 94.7%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
|||||:|||||

```
Db      417 CTTTACTTCATGCTCTTTG 399

RESULT 46
US-10-301-480-1169441/c
; Sequence 1169441, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1169441
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1169441

Query Match      91.6%; Score 17.4; DB 12; Length 997;
Best Local Similarity 94.7%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTTTACTTCATGCTCTTTG 19
      |||||
Db      417 CTTTACTTCATGCTCTTTG 399

RESULT 47
US-10-437-963-53068/c
; Sequence 53068, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 53068
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55303C.1
US-10-437-963-53068

Query Match      91.6%; Score 17.4; DB 8; Length 1371;
Best Local Similarity 94.7%; Pred. No. 5.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTTTACTTCATGCTCTTTG 19
      |||||
Db      135 CTTTGCTTCATGCTCTTTG 117

RESULT 48
US-09-946-290-7
; Sequence 7, Application US/09946290
; Publication No. US20030211475A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, Christopher J.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING PATHWAY-SPECIFIC REPORTERS AND
; FILE REFERENCE: 9301-040
; CURRENT APPLICATION NUMBER: US/09/946,290
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/282,243
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1001)..(1522)
US-09-946-290-7

Query Match      91.6%; Score 17.4; DB 3; Length 1725;
Best Local Similarity 94.7%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTTTACTTCATGCTCTTTG 19
      |||||
Db      386 CTTTACTTCATGCTCTTTG 404

RESULT 49
US-10-505-486-196/c
; Sequence 196, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 196
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Human
US-10-505-486-196

Query Match      91.6%; Score 17.4; DB 10; Length 1788;
Best Local Similarity 94.7%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTTTACTTCATGCTCTTTG 19
      |||||
Db      1059 CTTTACTTCATGCTCTTTG 1041

RESULT 50
US-11-116-881A-1152
; Sequence 1152, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
```

```
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1152
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-11-116-881A-1152

Query Match      86.3%; Score 16.4; DB 16; Length 168;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      35 TTTACTTCTTAGTCTTTG 52

RESULT 51
US-10-301-480-242885
; Sequence 242885, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242885
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-242885

Query Match      86.3%; Score 16.4; DB 12; Length 366;
Best Local Similarity 94.4%; Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      44 TTTACTTCAAAAGTCTTTG 61

RESULT 52
US-10-301-480-856294
; Sequence 856294, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 856294
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-856294

Query Match      86.3%; Score 16.4; DB 12; Length 366;
Best Local Similarity 94.4%; Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      44 TTTACTTCAAAAGTCTTTG 61

RESULT 53
US-09-925-065A-148754
; Sequence 148754, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148754
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-148754

Query Match      86.3%; Score 16.4; DB 4; Length 367;
Best Local Similarity 94.4%; Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTTG 19
Db      44 TTTACTTCAAAAGTCTTTG 61

RESULT 54
US-09-925-065A-148754
; Sequence 148754, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148754
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-148754

Query Match 86.3%; Score 16.4; DB 5; Length 367;
Best Local Similarity 94.4%; Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGCTTTG 19
Db 44 TTTACTTCAAAGTCTTTG 61

RESULT 55

US-09-925-065A-500185
; Sequence 500185, Application US/09925065A
; Publication No. US20040181049A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 500185
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-500185

Query Match 86.3%; Score 16.4; DB 4; Length 572;
Best Local Similarity 94.4%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTT 18
Db 230 CTTCACTTCATAGCTTTT 247

RESULT 56

US-09-925-065A-500185

; Sequence 500185, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 500185
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-500185

Query Match 86.3%; Score 16.4; DB 5; Length 572;
Best Local Similarity 94.4%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTT 18
Db 230 CTTCACTTCATAGCTTTT 247

RESULT 57

US-10-972-079-81269
; Sequence 81269, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE OF INVENTION: LIVESTOCK
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81269
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Chicken 19866894374567_3
US-10-972-079-81269

Query Match 86.3%; Score 16.4; DB 10; Length 576;
Best Local Similarity 94.4%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGCTTTG 19
Db 528 TTTACTTCTTAGTCTTTG 545

RESULT 58

US-09-925-065A-951021

; Sequence 951021, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 951021
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-951021

Query Match 86.3%; Score 16.4; DB 4; Length 595;
Best Local Similarity 94.4%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTT 18
||| |||||
Db 233 CTTCACTTCATAGTCCTTT 250

RESULT 59
US-09-925-065A-951021
; Sequence 951021, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 951021
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-951021

Query Match 86.3%; Score 16.4; DB 5; Length 595;
Best Local Similarity 94.4%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTT 18
||| |||||
Db 233 CTTCACTTCATAGTCCTTT 250

RESULT 60
US-10-425-115-71652
; Sequence 71652, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 71652
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165350C.1
US-10-425-115-71652

Query Match 86.3%; Score 16.4; DB 9; Length 602;
Best Local Similarity 94.4%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTT 18
||| |||||
Db 363 CTTCACTTCATAGTCCTTT 380

RESULT 61
US-09-925-065A-688576/c
; Sequence 688576, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 688576
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-688576

Query Match 86.3%; Score 16.4; DB 4; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTT 18
||| |||||
Db 476 CTTTCTTCATAGTCCTTT 459

RESULT 62
US-09-925-065A-688577/c

```
/ Sequence 688577, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 688577
/ LENGTH: 680
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-688577
```

```
Query Match      86.3%; Score 16.4; DB 4; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 CTTTACTTCATAGCTTTT 18
      ||||| ||||| ||||| |||||
Db      476 CTTTCTTCATAGCTTTT 459
```

```
RESULT 63
US-09-925-065A-688578/c
/ Sequence 688578, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 688578
/ LENGTH: 680
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-688578
```

```
Query Match      86.3%; Score 16.4; DB 4; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 CTTTACTTCATAGCTTTT 18
      ||||| ||||| ||||| |||||
Db      476 CTTTCTTCATAGCTTTT 459
```

```
RESULT 64
US-09-925-065A-688579/c
/ Sequence 688579, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 688579
/ LENGTH: 680
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-688579
```

```
Query Match      86.3%; Score 16.4; DB 4; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 CTTTACTTCATAGCTTTT 18
      ||||| ||||| ||||| |||||
Db      476 CTTTCTTCATAGCTTTT 459
```

```
RESULT 65
US-09-925-065A-688580/c
/ Sequence 688580, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 688580
/ LENGTH: 680
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-688580
```

```
Query Match      86.3%; Score 16.4; DB 4; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 CTTTACTTCATAGCTTTT 18
```



```
Db          476 CTTTCTTCATAGTCTTT 459
||||| ||||| ||||| ||||| |||||
Query Match      86.3%; Score 16.4; DB 5; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 CTTTACTTCATAGTCTTT 18
||||| ||||| ||||| ||||| |||||
Db          476 CTTTCTTCATAGTCTTT 459

RESULT 66
US-09-925-065A-688576/c
; Sequence 688576, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 688576
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-688576

Query Match      86.3%; Score 16.4; DB 5; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 CTTTACTTCATAGTCTTT 18
||||| ||||| ||||| ||||| |||||
Db          476 CTTTCTTCATAGTCTTT 459

RESULT 67
US-09-925-065A-688577/c
; Sequence 688577, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 688577
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-688577

Query Match      86.3%; Score 16.4; DB 5; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 CTTTACTTCATAGTCTTT 18
||||| ||||| ||||| ||||| |||||
Db          476 CTTTCTTCATAGTCTTT 459

RESULT 68
US-09-925-065A-688578/c
; Sequence 688578, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 688578
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-688578

Query Match      86.3%; Score 16.4; DB 5; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 CTTTACTTCATAGTCTTT 18
||||| ||||| ||||| ||||| |||||
Db          476 CTTTCTTCATAGTCTTT 459

RESULT 69
US-09-925-065A-688579/c
; Sequence 688579, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 688579
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-688579
```

```
Query Match      86.3%; Score 16.4; DB 5; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTT 18
    ||||| ||||| ||||| |||||
Db 476 CTTTCTTCATAGTCCTTT 459

RESULT 70
US-09-925-065A-688580/c
; Sequence 688580, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 688580
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-688580

Query Match      86.3%; Score 16.4; DB 5; Length 680;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTT 18
    ||||| ||||| ||||| |||||
Db 476 CTTTCTTCATAGTCCTTT 459

RESULT 71
US-09-925-065A-75210/c
; Sequence 75210, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75210
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-75210/c

Query Match      86.3%; Score 16.4; DB 4; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 201 TTTTATTCATAGTCCTTTG 184

RESULT 72
US-09-925-065A-75211/c
; Sequence 75211, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75211
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-75211

Query Match      86.3%; Score 16.4; DB 4; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 201 TTTTATTCATAGTCCTTTG 184

RESULT 73
US-09-925-065A-75210/c
; Sequence 75210, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75210
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-75210/c
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75210
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-75210

Query Match      86.3%; Score 16.4; DB 5; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 TTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db  201 TTTATTTTCATAGTCTTTG 184

RESULT 74
US-09-925-065A-75211/c
; Sequence 75211, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75211
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-75211

Query Match      86.3%; Score 16.4; DB 5; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 TTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db  201 TTTATTTTCATAGTCTTTG 184

RESULT 75
US-10-301-480-176449/c
; Sequence 176449, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 176449
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-176449

Query Match      86.3%; Score 16.4; DB 12; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 TTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db  201 TTTATTTTCATAGTCTTTG 184

RESULT 76
US-10-301-480-176450/c
; Sequence 176450, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 176450
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-176450

Query Match      86.3%; Score 16.4; DB 12; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 TTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db  201 TTTATTTTCATAGTCTTTG 184

RESULT 77
US-10-301-480-789858/c
; Sequence 789858, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 789858
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-789858

Query Match      86.3%; Score 16.4; DB 12; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 TTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db  201 TTTATTTTCATAGTCTTTG 184
```

```
Db      201 TTTATTTTCATAGTCTTTG 184
|||||
RESULT 78
US-10-301-480-789859/c
; Sequence 789859, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301.480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/261,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 789859
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-789859
Query Match      86.3%; Score 16.4; DB 12; Length 708;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      2 TTTACTTCATAGTCTTTG 19
|||||
Db      201 TTTATTTTCATAGTCTTTG 184
|||||
RESULT 79
US-09-925-065A-549225/c
; Sequence 549225, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 549225
; LENGTH: 802
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-925-065A-549225
Query Match      86.3%; Score 16.4; DB 4; Length 802;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      2 TTTACTTCATAGTCTTTG 19
|||||
Db      744 TTTACTTCAAAGTCTTTG 727
|||||
RESULT 80
US-09-925-065A-549225/c
; Sequence 549225, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 549225
; LENGTH: 802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-549225
Query Match      86.3%; Score 16.4; DB 5; Length 802;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      2 TTTACTTCATAGTCTTTG 19
|||||
Db      744 TTTACTTCAAAGTCTTTG 727
|||||
RESULT 81
US-10-301-480-527175/c
; Sequence 527175, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 527175
; LENGTH: 802
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-527175
Query Match      86.3%; Score 16.4; DB 12; Length 802;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      2 TTTACTTCATAGTCTTTG 19
|||||
Db      744 TTTACTTCAAAGTCTTTG 727
|||||
RESULT 82
US-10-301-480-1140584/c
```

; Sequence 1140584, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1140584
; LENGTH: 802
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1140584

Query Match 86.3%; Score 16.4; DB 12; Length 802;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTTG 19
|||||
Db 744 TTTACTCAAGCTTTG 727

RESULT 83
US-10-424-599-43257
; Sequence 43257, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43257
; LENGTH: 895
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13905C.1
US-10-424-599-43257

Query Match 86.3%; Score 16.4; DB 8; Length 895;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 18
|||||
Db 124 CTTTCTTCATAGTCTTT 141

RESULT 84
US-10-437-963-65932
; Sequence 65932, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 65932
; LENGTH: 995
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66931C.1
US-10-437-963-65932

Query Match 86.3%; Score 16.4; DB 8; Length 995;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 18
|||||
Db 439 CTTTGCTTCATAGTCTTT 456

RESULT 85
US-10-301-480-584099
; Sequence 584099, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 584099
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-584099

Query Match 86.3%; Score 16.4; DB 12; Length 999;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTTG 19
|||||
Db 597 TTCACCTTCATAGTCTTTG 614

RESULT 86
US-10-301-480-1197508
; Sequence 1197508, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1197508
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1197508

Query Match      86.3%; Score 16.4; DB 12; Length 999;
Best Local Similarity 94.4%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTTG 19
   |||||||
Db 597 TTCACCTTCATAGTCTTTG 614

RESULT 87
US-09-925-065A-49510/c
; Sequence 49510, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49510
; LENGTH: 2548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-49510

Query Match      86.3%; Score 16.4; DB 5; Length 2548;
Best Local Similarity 94.4%; Pred. No. 1.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTTG 19
   |||||||
Db 408 TTTACTTCATATCTTTG 391

RESULT 89
US-10-301-480-150748/c
; Sequence 150748, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150748
; LENGTH: 2548
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-150748

Query Match      86.3%; Score 16.4; DB 12; Length 2548;
Best Local Similarity 94.4%; Pred. No. 1.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTTTG 19
   |||||||
Db 408 TTTACTTCATATCTTTG 391

RESULT 90
US-10-301-480-764157/c
; Sequence 764157, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 764157
; LENGTH: 2548
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-764157
```

Query Match 86.3%; Score 16.4; DB 12; Length 2548;
Best Local Similarity 94.4%; Pred. No. 1.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19
| | | | | | | | | | | | | | | | | | | | | |
Db 408 TTTACTTCATATTCCTTTG 391

RESULT 91

US-10-750-185-53364/c
; Sequence 53364, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53364
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Bovine 198668808989808
US-10-750-185-53364

Query Match 86.3%; Score 16.4; DB 10; Length 3106;
Best Local Similarity 94.4%; Pred. No. 1.9e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTTG 19
| | | | | | | | | | | | | | | | | | | | | |
Db 305 TCTACTTCATAGTCTTTG 288

RESULT 92

US-10-750-623-53364/c
; Sequence 53364, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53364
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Bovine 198668808989808
US-10-750-623-53364

Query Match 86.3%; Score 16.4; DB 10; Length 3106;
Best Local Similarity 94.4%; Pred. No. 1.9e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 TTTACTTCATAGTCTTTG 19
| | | | | | | | | | | | | | | | | | | | | |
Db 305 TCTACTTCATAGTCTTTG 288

RESULT 93

US-11-083-784-470117/c
; Sequence 470117, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 470117
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-083-784-470117

Query Match 84.2%; Score 16; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTT 17
| | | | | | | | | | | | | | | | | | | | | |
Db 18 TTTACTTCATAGTCTT 3

RESULT 94

US-11-083-784-470217/c
; Sequence 470217, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 470217
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-083-784-470217

Query Match 84.2%; Score 16; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTT 17
|||||
Db 18 TTTACTTCATAGTCTT 3

RESULT 95
US-11-101-244-470117/c
; Sequence 470117, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 470117
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-470117

Query Match 84.2%; Score 16; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTT 17
|||||
Db 18 TTTACTTCATAGTCTT 3

RESULT 96
US-11-101-244-470217/c
; Sequence 470217, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 470217
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-470217

Query Match 84.2%; Score 16; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCTT 17
|||||
Db 18 TTTACTTCATAGTCTT 3

RESULT 97
US-11-060-756-43358/c
; Sequence 43358, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43358
; LENGTH: 25
; TYPE: DNA
; ORGANISM: probe
US-11-060-756-43358

Query Match 84.2%; Score 16; DB 13; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCT 16
|||||
Db 16 CTTTACTTCATAGTCT 1

RESULT 98
US-11-060-756-43361/c
; Sequence 43361, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43361
; LENGTH: 25
; TYPE: DNA
; ORGANISM: probe
US-11-060-756-43361

Query Match 84.2%; Score 16; DB 13; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TACTTCATAGTCTTTG 19
|||||
Db 25 TACTTCATAGTCTTTG 10

RESULT 99
US-11-175-859-39398/c
; Sequence 39398, Application US/11175859
; Publication No. US20060024715A1
; GENERAL INFORMATION:
; APPLICANT: Affymetrix, Inc.

```
; APPLICANT: Liu, Guoying et al.
; TITLE OF INVENTION: Method of Analysis of Human Polymorphism
; FILE REFERENCE: 3650.1
; CURRENT APPLICATION NUMBER: US/11/175,859
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: US 60/585,352
; PRIOR FILING DATE: 2004-07-02
; NUMBER OF SEQ ID NOS: 116251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39398
; LENGTH: 50
; TYPE: DNA
; ORGANISM: homo sapien
US-11-175-859-39398
```

```
Query Match      84.2%; Score 16; DB 16; Length 50;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 CTTTACTTCATAGTCTTT 18
      |||:|||||
Db      30 CATTCCTTCATAGTCTTT 13
```

```
RESULT 100
US-10-301-480-242884
; Sequence 242884, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 242884
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-242884
```

```
Query Match      84.2%; Score 16; DB 12; Length 366;
Best Local Similarity 88.9%; Pred. No. 2.2e+03;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 TTTACTTCATAGTCTTTG 19
      |||:|||||
Db      44 TTTAYTTCAAAGTCTTTG 61
```

```
Search completed: June 10, 2006, 20:51:02
Job time : 744.7 secs
```

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:34:22 ; Search time 9.18333 seconds
(without alignments)
262.600 Million cell updates/sec

Title: US-10-600-816-32

Perfect score: 19

Sequence: 1 cttactcatgctcttg 19

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 253354 seqs, 63461778 residues

Total number of hits satisfying chosen parameters: 506708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database : Published Applications NA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	15.4	81.1	522	6	US-10-471-571A-1359
C 2	15.4	81.1	2163	7	US-11-217-529-78062
C 3	15.4	81.1	5376	7	US-11-313-450-17
C 4	15	78.9	1446	6	US-10-953-349-5467
C 5	14.8	77.9	588	7	US-11-183-218-74
C 6	14.8	77.9	654	7	US-11-217-529-76788
C 7	14.8	77.9	676	6	US-10-953-349-6204
C 8	14.8	77.9	956	6	US-10-953-349-2872
C 9	14.8	77.9	1212	7	US-11-217-529-78358
C 10	14.8	77.9	1432	6	US-10-953-349-2801
C 11	14.8	77.9	1882	6	US-10-541-947-5
C 12	14.8	77.9	3551	7	US-11-293-697-610
C 13	14.8	77.9	3816	7	US-11-293-697-662
C 14	14.8	77.9	4218	7	US-11-217-529-75740
C 15	14.4	75.8	225	7	US-11-217-529-11133
C 16	14.4	75.8	285	7	US-11-217-529-80640
C 17	14.4	75.8	432	7	US-11-217-529-475
C 18	14.4	75.8	1457	6	US-10-953-349-11495
C 19	14.4	75.8	3689	7	US-11-293-697-523
C 20	14.2	74.7	25	7	US-11-217-529-161119
C 21	14.2	74.7	25	7	US-11-217-529-172215
C 22	14.2	74.7	192	7	US-11-217-529-166960
C 23	14.2	74.7	255	7	US-11-217-529-80554
C 24	14.2	74.7	517	6	US-10-953-349-10182
C 25	14.2	74.7	606	7	US-11-217-529-82353
C 26	14.2	74.7	771	7	US-11-217-529-1071
C 27	14.2	74.7	771	7	US-11-217-529-76887
C 28	14.2	74.7	819	7	US-11-217-529-79869
C 29	14.2	74.7	912	6	US-10-511-937-501
C 30	14.2	74.7	936	6	US-10-953-349-5011
C 31	14.2	74.7	1333	6	US-10-953-349-7897
C 32	14.2	74.7	1333	7	US-11-285-701-17
C 33	14.2	74.7	1566	7	US-11-217-529-780
C 34	14.2	74.7	1572	7	US-11-217-529-77811
C 35	14.2	74.7	1860	7	US-11-293-697-1270
C 36	14.2	74.7	1910	7	US-11-293-697-155
C 37	14.2	74.7	1986	6	US-10-953-349-11687
C 38	14.2	74.7	2010	7	US-11-217-529-2941
C 39	14.2	74.7	2594	7	US-11-293-697-1439
C 40	14.2	74.7	2628	7	US-11-217-529-78237
C 41	14.2	74.7	4675	7	US-11-327-357-1
C 42	14	73.7	936	7	US-11-217-529-138
C 43	14	73.7	1422	6	US-10-953-349-2188
C 44	14	73.7	2013	6	US-10-953-349-8157
C 45	13.8	72.6	25	7	US-11-217-529-190653
C 46	13.8	72.6	183	6	US-10-471-571A-751
C 47	13.8	72.6	272	6	US-10-473-173-164
C 48	13.8	72.6	330	7	US-11-301-554-1379
C 49	13.8	72.6	378	7	US-11-217-529-174589
C 50	13.8	72.6	564	6	US-10-488-619-1992
C 51	13.8	72.6	696	7	US-11-217-529-82746
C 52	13.8	72.6	729	7	US-11-217-529-77789
C 53	13.8	72.6	841	6	US-10-953-349-8745
C 54	13.8	72.6	867	7	US-11-217-529-166678
C 55	13.8	72.6	885	6	US-10-471-571A-3451
C 56	13.8	72.6	900	6	US-10-471-571A-3035
C 57	13.8	72.6	946	6	US-10-953-349-7110
C 58	13.8	72.6	1017	6	US-10-471-571A-3891
C 59	13.8	72.6	1035	6	US-11-217-529-78906
C 60	13.8	72.6	1109	6	US-10-953-349-14861
C 61	13.8	72.6	1172	6	US-10-511-937-441
C 62	13.8	72.6	1191	6	US-10-471-571A-785
C 63	13.8	72.6	1240	6	US-10-953-349-20194
C 64	13.8	72.6	1245	7	US-11-217-529-78180
C 65	13.8	72.6	1272	6	US-10-953-349-1160
C 66	13.8	72.6	1315	6	US-10-953-349-20559
C 67	13.8	72.6	1332	6	US-10-953-349-15758
C 68	13.8	72.6	1333	6	US-10-953-349-17823
C 69	13.8	72.6	1341	6	US-10-471-571A-4559
C 70	13.8	72.6	1404	6	US-10-953-349-24131
C 71	13.8	72.6	1420	6	US-10-953-349-7778
C 72	13.8	72.6	1461	7	US-11-217-529-82109
C 73	13.8	72.6	1485	7	US-11-217-529-5036
C 74	13.8	72.6	1605	7	US-11-217-529-82729
C 75	13.8	72.6	1704	7	US-11-217-529-421
C 76	13.8	72.6	1800	7	US-11-217-529-4484
C 77	13.8	72.6	1806	6	US-10-953-349-20342
C 78	13.8	72.6	1843	6	US-10-953-349-1741
C 79	13.8	72.6	1854	7	US-11-217-529-534
C 80	13.8	72.6	1861	6	US-10-953-349-18787
C 81	13.8	72.6	1899	6	US-10-471-571A-4139
C 82	13.8	72.6	2004	7	US-11-217-529-75539
C 83	13.8	72.6	2031	7	US-11-217-529-565
C 84	13.8	72.6	2190	6	US-10-471-571A-1409
C 85	13.8	72.6	2320	6	US-10-953-349-8153
C 86	13.8	72.6	2577	7	US-11-293-697-52
C 87	13.8	72.6	2650	7	US-11-293-697-749
C 88	13.8	72.6	2854	7	US-11-293-697-2321
C 89	13.8	72.6	3102	7	US-11-217-529-82302
C 90	13.8	72.6	3273	7	US-11-217-529-4879
C 91	13.8	72.6	3447	7	US-11-217-529-77385
C 92	13.8	72.6	3528	7	US-11-217-529-3930
C 93	13.8	72.6	4746	6	US-10-953-349-2088
C 94	13.8	72.6	5283	7	US-11-217-529-5476
C 95	13.8	72.6	394191	6	US-10-506-549-3
C 96	13.4	70.5	25	7	US-11-217-529-74966
C 97	13.4	70.5	552	6	US-10-471-571A-1517
C 98	13.4	70.5	878	6	US-10-953-349-15823

C 99	13.4	70.5	966	7	US-11-217-529-269	Sequence 269, App	C 172	13.2	69.5	1605	7	US-11-217-529-75988	Sequence 75988, A
C 100	13.4	70.5	1023	7	US-11-253-300-6	Sequence 6, Appl1	C 173	13.2	69.5	1638	7	US-11-217-529-79707	Sequence 79707, A
C 101	13.4	70.5	1335	7	US-11-264-737-11	Sequence 11, Appl	C 174	13.2	69.5	1659	7	US-11-217-529-82648	Sequence 82648, A
C 102	13.4	70.5	1335	7	US-11-264-737-11	Sequence 11, Appl	C 175	13.2	69.5	1702	6	US-10-511-937-429	Sequence 429, App
C 103	13.4	70.5	1335	7	US-11-265-761-11	Sequence 11, Appl	C 176	13.2	69.5	1702	6	US-10-511-937-2866	Sequence 2866, App
C 104	13.4	70.5	1331	6	US-10-471-571A-2235	Sequence 2235, App	C 177	13.2	69.5	1713	7	US-11-217-529-75972	Sequence 75972, A
C 105	13.4	70.5	1384	6	US-10-953-349-15710	Sequence 15710, A	C 178	13.2	69.5	1776	7	US-11-217-529-1921	Sequence 1921, App
C 106	13.4	70.5	1774	7	US-11-293-697-2118	Sequence 2118, App	C 179	13.2	69.5	1796	6	US-10-953-349-20551	Sequence 20551, A
C 107	13.4	70.5	1837	6	US-10-953-349-16888	Sequence 16888, A	C 180	13.2	69.5	1833	7	US-11-217-529-80544	Sequence 80544, A
C 108	13.4	70.5	1979	6	US-10-953-349-860	Sequence 860, App	C 181	13.2	69.5	1837	6	US-10-953-349-18744	Sequence 18744, A
C 109	13.4	70.5	2163	7	US-11-217-529-6203	Sequence 6203, App	C 182	13.2	69.5	1875	7	US-11-217-529-4544	Sequence 4544, App
C 110	13.4	70.5	2298	6	US-10-505-928-487	Sequence 487, App	C 183	13.2	69.5	1905	7	US-11-217-529-82415	Sequence 82415, A
C 111	13.4	70.5	4131	7	US-11-217-529-158	Sequence 158, App	C 184	13.2	69.5	1939	7	US-11-293-697-1971	Sequence 1971, App
C 112	13.4	70.5	4524	7	US-11-217-529-80351	Sequence 80351, A	C 185	13.2	69.5	1993	7	US-11-293-697-392	Sequence 392, App
C 113	13.4	70.5	261789	7	US-11-260-842-1	Sequence 1, Appl1	C 186	13.2	69.5	2006	6	US-10-953-349-16902	Sequence 16902, A
C 114	13.4	70.5	261789	7	US-11-260-842-1	Sequence 1, Appl1	C 187	13.2	69.5	2106	7	US-11-217-529-77738	Sequence 77738, A
C 115	13.2	69.5	25	7	US-11-217-529-158697	Sequence 158697, A	C 188	13.2	69.5	2114	6	US-10-953-349-19677	Sequence 19677, A
C 116	13.2	69.5	51	7	US-11-143-642-938	Sequence 938, App	C 189	13.2	69.5	2140	7	US-11-293-697-271	Sequence 271, App
C 117	13.2	69.5	51	7	US-11-245-448-519	Sequence 519, App	C 190	13.2	69.5	2142	7	US-11-217-529-75343	Sequence 75343, A
C 118	13.2	69.5	165	6	US-10-471-571A-3685	Sequence 3685, App	C 191	13.2	69.5	2228	7	US-11-293-697-744	Sequence 744, App
C 119	13.2	69.5	207	6	US-10-560-723-31	Sequence 31, Appl	C 192	13.2	69.5	2252	7	US-11-217-529-78725	Sequence 78725, A
C 120	13.2	69.5	327	7	US-11-217-529-173920	Sequence 173920, A	C 193	13.2	69.5	2298	7	US-11-217-529-3750	Sequence 3750, App
C 121	13.2	69.5	369	7	US-11-217-529-77256	Sequence 77256, A	C 194	13.2	69.5	2305	7	US-11-293-697-535	Sequence 535, App
C 122	13.2	69.5	489	7	US-11-217-529-75638	Sequence 75638, A	C 195	13.2	69.5	2340	6	US-10-196-749-379	Sequence 379, App
C 123	13.2	69.5	543	7	US-11-217-529-81867	Sequence 81867, A	C 196	13.2	69.5	2352	7	US-11-217-529-2563	Sequence 2563, App
C 124	13.2	69.5	639	7	US-11-301-554-716	Sequence 716, App	C 197	13.2	69.5	2481	7	US-11-217-529-79177	Sequence 79177, A
C 125	13.2	69.5	672	7	US-11-217-529-2034	Sequence 2034, App	C 198	13.2	69.5	2682	7	US-11-217-529-5575	Sequence 5575, App
C 126	13.2	69.5	693	7	US-11-217-529-81027	Sequence 81027, A	C 199	13.2	69.5	2738	7	US-11-293-697-270	Sequence 270, App
C 127	13.2	69.5	699	6	US-10-471-571A-1461	Sequence 1461, App	C 200	13.2	69.5	2859	7	US-11-217-529-4980	Sequence 4980, App
C 128	13.2	69.5	702	6	US-10-471-571A-1809	Sequence 1809, App	C 201	13.2	69.5	2910	7	US-11-217-529-77545	Sequence 77545, A
C 129	13.2	69.5	741	6	US-10-471-571A-959	Sequence 959, App	C 202	13.2	69.5	3038	6	US-10-196-749-69	Sequence 69, Appl
C 130	13.2	69.5	765	6	US-10-488-619-2769	Sequence 2769, App	C 203	13.2	69.5	3219	7	US-11-217-529-79	Sequence 79, Appl
C 131	13.2	69.5	779	6	US-10-953-349-6577	Sequence 6577, App	C 204	13.2	69.5	3373	7	US-11-293-697-1419	Sequence 1419, App
C 132	13.2	69.5	807	7	US-11-217-529-2399	Sequence 2399, App	C 205	13.2	69.5	3617	6	US-10-982-908-13	Sequence 13, Appl
C 133	13.2	69.5	810	6	US-10-953-349-39311	Sequence 39311, A	C 206	13.2	69.5	3628	7	US-11-293-697-353	Sequence 353, App
C 134	13.2	69.5	837	7	US-11-217-529-166621	Sequence 166621, A	C 207	13.2	69.5	3741	6	US-10-505-928-289	Sequence 289, App
C 135	13.2	69.5	868	6	US-10-953-349-512	Sequence 512, App	C 208	13.2	69.5	3804	7	US-11-217-529-78013	Sequence 78013, A
C 136	13.2	69.5	884	6	US-10-953-349-12775	Sequence 12775, A	C 209	13.2	69.5	4899	7	US-11-217-529-4664	Sequence 4664, App
C 137	13.2	69.5	891	7	US-11-217-529-3027	Sequence 3027, App	C 210	13.2	69.5	4995	7	US-11-217-529-78781	Sequence 78781, A
C 138	13.2	69.5	891	7	US-11-217-529-3027	Sequence 3027, App	C 211	13.2	69.5	5130	7	US-11-217-529-80522	Sequence 80522, A
C 139	13.2	69.5	900	6	US-10-471-571A-2405	Sequence 2405, App	C 212	13.2	69.5	5193	7	US-11-317-983-3	Sequence 3, Appl1
C 140	13.2	69.5	946	6	US-10-953-349-660	Sequence 660, App	C 213	13.2	69.5	5339	7	US-11-317-983-24	Sequence 24, Appl1
C 141	13.2	69.5	952	6	US-10-953-349-6599	Sequence 6599, App	C 214	13.2	69.5	10211	6	US-10-505-928-326	Sequence 326, App
C 142	13.2	69.5	957	6	US-10-471-571A-1663	Sequence 1663, App	C 215	13.2	69.5	10211	6	US-11-217-529-125975	Sequence 125975, A
C 143	13.2	69.5	1000	7	US-11-233-726-2	Sequence 2, Appl1	C 216	13.2	69.5	10211	6	US-11-217-529-80266	Sequence 80266, A
C 144	13.2	69.5	1024	6	US-10-953-349-23848	Sequence 23848, A	C 217	13.2	69.5	10211	6	US-11-217-529-79207	Sequence 79207, A
C 145	13.2	69.5	1029	6	US-10-953-349-30803	Sequence 30803, A	C 218	13.2	69.5	1077	7	US-11-217-529-79208	Sequence 79208, A
C 146	13.2	69.5	1077	6	US-10-953-349-2626	Sequence 2626, App	C 219	13.2	69.5	1497	7	US-11-217-529-2278	Sequence 2278, App
C 147	13.2	69.5	1083	6	US-10-953-349-23918	Sequence 23918, A	C 220	13.2	69.5	1618	6	US-10-953-349-6398	Sequence 6398, App
C 148	13.2	69.5	1097	6	US-10-953-349-16654	Sequence 16654, A	C 221	13.2	69.5	1719	7	US-11-217-529-79365	Sequence 79365, A
C 149	13.2	69.5	1137	7	US-11-217-529-1773	Sequence 1773, App	C 222	13.2	69.5	1920	7	US-11-217-529-90	Sequence 90, Appl
C 150	13.2	69.5	1137	7	US-11-217-529-166563	Sequence 166563, A	C 223	12.8	67.4	25	7	US-11-217-529-156849	Sequence 156849, A
C 151	13.2	69.5	1143	7	US-11-217-529-4165	Sequence 4165, App	C 224	12.8	67.4	162	7	US-11-217-529-82716	Sequence 82716, A
C 152	13.2	69.5	1162	6	US-10-953-349-22893	Sequence 22893, A	C 225	12.8	67.4	300	6	US-10-471-571A-1595	Sequence 1595, App
C 153	13.2	69.5	1254	6	US-10-953-349-5222	Sequence 5222, App	C 226	12.8	67.4	326	7	US-11-301-554-1306	Sequence 1306, App
C 154	13.2	69.5	1260	7	US-11-256-428-66	Sequence 66, Appl	C 227	12.8	67.4	333	7	US-11-301-554-488	Sequence 488, App
C 155	13.2	69.5	1263	6	US-10-953-349-15750	Sequence 15750, A	C 228	12.8	67.4	393	7	US-11-301-554-778	Sequence 778, App
C 156	13.2	69.5	1281	6	US-10-953-349-10590	Sequence 10590, A	C 229	12.8	67.4	397	7	US-11-301-554-694	Sequence 694, App
C 157	13.2	69.5	1297	6	US-10-953-349-22597	Sequence 22597, A	C 230	12.8	67.4	414	6	US-10-471-571A-3571	Sequence 3571, App
C 158	13.2	69.5	1297	6	US-10-953-349-11295	Sequence 11295, A	C 231	12.8	67.4	423	7	US-11-217-529-81957	Sequence 81957, A
C 159	13.2	69.5	1300	6	US-10-953-349-19528	Sequence 19528, A	C 232	12.8	67.4	446	7	US-11-217-529-81952	Sequence 81952, A
C 160	13.2	69.5	1333	6	US-10-953-349-7897	Sequence 7897, App	C 233	12.8	67.4	446	6	US-10-953-349-36171	Sequence 36171, A
C 161	13.2	69.5	1333	7	US-11-285-701-17	Sequence 17, Appl	C 234	12.8	67.4	480	7	US-11-217-529-81535	Sequence 81535, A
C 162	13.2	69.5	1341	7	US-11-217-529-82122	Sequence 82122, A	C 235	12.8	67.4	542	6	US-10-488-619-1146	Sequence 1146, App
C 163	13.2	69.5	1349	6	US-10-953-349-15273	Sequence 15273, A	C 236	12.8	67.4	564	7	US-11-217-529-5892	Sequence 5892, App
C 164	13.2	69.5	1352	6	US-10-953-349-20777	Sequence 20777, A	C 237	12.8	67.4	618	6	US-10-953-349-9081	Sequence 9081, App
C 165	13.2	69.5	1358	6	US-10-953-349-33616	Sequence 33616, A	C 238	12.8	67.4	715	6	US-10-953-349-6110	Sequence 6110, App
C 166	13.2	69.5	1380	6	US-10-471-571A-1443	Sequence 1443, App	C 239	12.8	67.4	750	6	US-10-488-619-1179	Sequence 1179, App
C 167	13.2	69.5	1383	7	US-11-217-529-296	Sequence 296, App	C 240	12.8	67.4	753	6	US-10-953-349-9366	Sequence 9366, App
C 168	13.2	69.5	1413	7	US-11-217-529-5563	Sequence 5563, App	C 241	12.8	67.4	855	7	US-11-217-529-78658	Sequence 78658, A
C 169	13.2	69.5	1446	7	US-11-217-529-1760	Sequence 1760, App	C 242	12.8	67.4	888	7	US-11-217-529-1118	Sequence 1118, App
C 170	13.2	69.5	1560	6	US-10-471-571A-5355	Sequence 5355, App	C 243	12.8	67.4	1003	7	US-11-226-605-53	Sequence 53, Appl
C 171	13.2	69.5	1573	6	US-10-953-349-38598	Sequence 38598, A	C 244	12.8	67.4	1041	7	US-11-217-529-79470	Sequence 79470, A

C 245	12.8	67.4	1059	7	US-11-217-529-82500	Sequence 82500, A	C 318	12.8	67.4	11978	7	US-11-257-851A-63	Sequence 63, Appl
C 246	12.8	67.4	1156	6	US-10-953-349-15088	Sequence 15088, A	C 319	12.6	66.3	25	7	US-11-217-529-88708	Sequence 88708, A
C 247	12.8	67.4	1172	6	US-10-953-349-4182	Sequence 4182, A	C 320	12.6	66.3	25	7	US-11-217-529-121670	Sequence 121670, A
C 248	12.8	67.4	1174	6	US-10-953-349-8797	Sequence 8797, A	C 321	12.6	66.3	159	6	US-10-471-571A-1507	Sequence 1507, Ap
C 249	12.8	67.4	1200	6	US-10-471-571A-4647	Sequence 4647, Ap	C 322	12.6	66.3	339	7	US-11-217-529-166820	Sequence 166820, A
C 250	12.8	67.4	1224	6	US-10-953-349-9377	Sequence 9377, Ap	C 323	12.6	66.3	342	7	US-11-217-529-78472	Sequence 78472, A
C 251	12.8	67.4	1224	6	US-11-217-529-79666	Sequence 79666, A	C 324	12.6	66.3	384	7	US-11-217-529-75777	Sequence 75777, A
C 252	12.8	67.4	1233	7	US-11-293-697-1566	Sequence 1566, Ap	C 325	12.6	66.3	411	7	US-11-217-529-5764	Sequence 5764, Ap
C 253	12.8	67.4	1249	6	US-10-953-349-14358	Sequence 14358, A	C 326	12.6	66.3	435	7	US-11-217-529-75873	Sequence 75873, A
C 254	12.8	67.4	1254	7	US-11-217-529-77045	Sequence 77045, A	C 327	12.6	66.3	441	7	US-11-217-529-2542	Sequence 2542, Ap
C 255	12.8	67.4	1300	6	US-10-953-349-14532	Sequence 14532, A	C 328	12.6	66.3	468	7	US-11-301-554-604	Sequence 604, App
C 256	12.8	67.4	1306	6	US-10-953-349-31482	Sequence 31482, A	C 329	12.6	66.3	486	7	US-11-217-529-5793	Sequence 5793, Ap
C 257	12.8	67.4	1326	6	US-10-953-349-17594	Sequence 17594, A	C 330	12.6	66.3	507	7	US-11-217-529-190968	Sequence 190968, A
C 258	12.8	67.4	1380	7	US-11-217-529-76526	Sequence 76526, A	C 331	12.6	66.3	510	7	US-11-217-529-3168	Sequence 3168, Ap
C 259	12.8	67.4	1380	7	US-11-217-529-78516	Sequence 78516, A	C 332	12.6	66.3	524	6	US-10-953-349-5669	Sequence 5669, Ap
C 260	12.8	67.4	1431	7	US-11-217-529-1404	Sequence 1404, Ap	C 333	12.6	66.3	536	6	US-10-488-619-1668	Sequence 1668, Ap
C 261	12.8	67.4	1449	7	US-11-217-529-78578	Sequence 78578, A	C 334	12.6	66.3	576	6	US-10-488-619-1882	Sequence 1882, Ap
C 262	12.8	67.4	1470	6	US-10-953-349-82094	Sequence 82094, A	C 335	12.6	66.3	591	7	US-11-217-529-174057	Sequence 174057, A
C 263	12.8	67.4	1470	7	US-11-217-529-3767	Sequence 3767, Ap	C 336	12.6	66.3	603	7	US-11-217-529-765	Sequence 765, App
C 264	12.8	67.4	1473	6	US-10-953-349-24770	Sequence 24770, A	C 337	12.6	66.3	615	7	US-11-217-529-1614	Sequence 1614, Ap
C 265	12.8	67.4	1488	6	US-10-526-429-39	Sequence 39, Appl	C 338	12.6	66.3	624	6	US-10-471-571A-5287	Sequence 5287, Ap
C 266	12.8	67.4	1515	7	US-11-217-529-4555	Sequence 4555, Ap	C 339	12.6	66.3	630	6	US-11-217-529-78824	Sequence 78824, A
C 267	12.8	67.4	1542	7	US-11-217-529-1892	Sequence 1892, Ap	C 340	12.6	66.3	644	6	US-10-196-749-213	Sequence 213, App
C 268	12.8	67.4	1554	7	US-11-217-529-3525	Sequence 3525, Ap	C 341	12.6	66.3	644	7	US-11-101-316-55	Sequence 55, Appl
C 269	12.8	67.4	1581	7	US-11-217-529-82085	Sequence 82085, A	C 342	12.6	66.3	654	7	US-11-217-529-2467	Sequence 2467, Ap
C 270	12.8	67.4	1641	6	US-11-217-529-78032	Sequence 78032, A	C 343	12.6	66.3	658	6	US-10-488-619-2511	Sequence 2511, Ap
C 271	12.8	67.4	1654	6	US-10-953-349-11766	Sequence 11766, A	C 344	12.6	66.3	695	6	US-10-488-619-1881	Sequence 1881, Ap
C 272	12.8	67.4	1653	7	US-11-217-529-3155	Sequence 3155, Ap	C 345	12.6	66.3	717	7	US-11-217-529-81169	Sequence 81169, A
C 273	12.8	67.4	1665	7	US-11-217-529-1913	Sequence 1913, Ap	C 346	12.6	66.3	731	7	US-11-217-529-166792	Sequence 166792, A
C 274	12.8	67.4	1689	7	US-11-217-529-80752	Sequence 80752, A	C 347	12.6	66.3	748	7	US-11-217-529-76004	Sequence 76004, A
C 275	12.8	67.4	1704	6	US-10-471-571A-917	Sequence 917, App	C 348	12.6	66.3	742	6	US-10-511-937-2887	Sequence 2887, Ap
C 276	12.8	67.4	1749	7	US-11-217-529-79614	Sequence 79614, A	C 349	12.6	66.3	752	6	US-10-953-349-16464	Sequence 16464, Ap
C 277	12.8	67.4	1761	7	US-11-217-529-78138	Sequence 78138, A	C 350	12.6	66.3	771	7	US-11-217-529-1659	Sequence 1659, Ap
C 278	12.8	67.4	1961	7	US-11-293-697-2247	Sequence 2247, Ap	C 351	12.6	66.3	785	6	US-10-953-349-18263	Sequence 18263, A
C 279	12.8	67.4	1983	6	US-10-953-349-5126	Sequence 5126, Ap	C 352	12.6	66.3	792	6	US-10-488-619-2510	Sequence 2510, Ap
C 280	12.8	67.4	2004	7	US-11-217-529-1382	Sequence 1382, Ap	C 353	12.6	66.3	816	6	US-10-953-349-11989	Sequence 11989, A
C 281	12.8	67.4	2031	7	US-11-217-529-78529	Sequence 78529, A	C 354	12.6	66.3	826	6	US-11-217-529-76874	Sequence 76874, A
C 282	12.8	67.4	2088	7	US-11-217-529-586	Sequence 586, App	C 355	12.6	66.3	820	6	US-10-953-349-12479	Sequence 12479, A
C 283	12.8	67.4	2145	7	US-11-217-529-80079	Sequence 80079, A	C 356	12.6	66.3	822	6	US-11-217-529-75872	Sequence 75872, A
C 284	12.8	67.4	2145	7	US-11-217-529-81000	Sequence 81000, A	C 357	12.6	66.3	909	6	US-10-953-349-1444	Sequence 1444, Ap
C 285	12.8	67.4	2158	7	US-11-293-697-2369	Sequence 2369, Ap	C 358	12.6	66.3	915	7	US-11-217-529-1621	Sequence 1621, Ap
C 286	12.8	67.4	2283	7	US-11-217-529-3378	Sequence 3378, Ap	C 359	12.6	66.3	915	7	US-11-217-529-191042	Sequence 191042, A
C 287	12.8	67.4	2298	6	US-10-511-937-525	Sequence 525, App	C 360	12.6	66.3	924	7	US-11-217-529-2811	Sequence 2811, Ap
C 288	12.8	67.4	2514	7	US-11-217-529-5277	Sequence 5277, Ap	C 361	12.6	66.3	924	7	US-11-217-529-173251	Sequence 173251, A
C 289	12.8	67.4	2517	7	US-11-293-697-1130	Sequence 1130, Ap	C 362	12.6	66.3	937	6	US-10-953-349-28474	Sequence 28474, A
C 290	12.8	67.4	2664	7	US-11-217-529-79013	Sequence 79013, A	C 363	12.6	66.3	954	7	US-11-217-529-78390	Sequence 78390, A
C 291	12.8	67.4	2979	7	US-11-253-453-4	Sequence 4, Appl	C 364	12.6	66.3	971	1	US-09-949-525-63	Sequence 63, Appl
C 292	12.8	67.4	2979	7	US-11-178-560-4	Sequence 4, Appl	C 365	12.6	66.3	997	6	US-10-953-349-19221	Sequence 19221, A
C 293	12.8	67.4	3027	6	US-10-471-571A-2401	Sequence 2401, Ap	C 366	12.6	66.3	1003	7	US-11-226-605-16	Sequence 16, Appl
C 294	12.8	67.4	3230	7	US-11-293-697-2392	Sequence 2392, Ap	C 367	12.6	66.3	1008	7	US-11-217-529-81788	Sequence 81788, A
C 295	12.8	67.4	3430	7	US-11-293-697-1442	Sequence 1442, Ap	C 368	12.6	66.3	1008	7	US-11-217-529-191099	Sequence 191099, A
C 296	12.8	67.4	3435	6	US-10-953-349-4556	Sequence 4556, Ap	C 369	12.6	66.3	1011	7	US-11-217-529-77524	Sequence 77524, A
C 297	12.8	67.4	3627	7	US-11-293-697-1570	Sequence 1570, Ap	C 370	12.6	66.3	1026	6	US-10-953-349-5364	Sequence 5364, Ap
C 298	12.8	67.4	3793	7	US-11-293-697-1352	Sequence 1352, Ap	C 371	12.6	66.3	1038	6	US-11-217-529-1880	Sequence 1880, Ap
C 299	12.8	67.4	3843	7	US-11-217-529-1951	Sequence 1951, Ap	C 372	12.6	66.3	1041	7	US-11-217-529-508	Sequence 508, App
C 300	12.8	67.4	3926	7	US-11-251-465-14	Sequence 14, Appl	C 373	12.6	66.3	1041	7	US-11-217-529-82578	Sequence 82578, A
C 301	12.8	67.4	4047	6	US-10-471-571A-3351	Sequence 3351, Ap	C 374	12.6	66.3	1050	6	US-10-953-349-10035	Sequence 10035, A
C 302	12.8	67.4	4321	7	US-11-313-836-7	Sequence 7, Appl	C 375	12.6	66.3	1104	7	US-11-217-529-82147	Sequence 82147, A
C 303	12.8	67.4	4357	6	US-10-511-937-510	Sequence 510, App	C 376	12.6	66.3	1112	6	US-10-953-349-25161	Sequence 25161, A
C 304	12.8	67.4	4434	7	US-11-217-529-76875	Sequence 76875, A	C 377	12.6	66.3	1128	7	US-11-217-529-3541	Sequence 3541, Ap
C 305	12.8	67.4	4473	7	US-11-217-529-1290	Sequence 1290, Ap	C 378	12.6	66.3	1131	7	US-11-217-529-77951	Sequence 77951, A
C 306	12.8	67.4	4576	6	US-10-531-965-1	Sequence 1, Appl	C 379	12.6	66.3	1141	6	US-10-953-349-18899	Sequence 18899, A
C 307	12.8	67.4	4611	7	US-11-217-529-191207	Sequence 191207, A	C 380	12.6	66.3	1141	6	US-10-953-349-22441	Sequence 22441, A
C 308	12.8	67.4	4901	6	US-10-953-349-1429	Sequence 1429, Ap	C 381	12.6	66.3	1149	7	US-11-217-529-76533	Sequence 76533, A
C 309	12.8	67.4	5047	6	US-11-217-529-575	Sequence 575, App	C 382	12.6	66.3	1150	6	US-10-953-349-23523	Sequence 23523, A
C 310	12.8	67.4	5493	7	US-11-217-529-91749	Sequence 91749, A	C 383	12.6	66.3	1161	6	US-10-511-937-3110	Sequence 3110, Ap
C 311	12.8	67.4	5734	6	US-10-505-928-28	Sequence 28, Appl	C 384	12.6	66.3	1161	6	US-10-953-349-4619	Sequence 4619, Ap
C 312	12.8	67.4	7884	7	US-11-217-529-3160	Sequence 3160, Ap	C 385	12.6	66.3	1184	6	US-10-953-349-26599	Sequence 26599, A
C 313	12.8	67.4	9737	1	US-09-484-331-22	Sequence 22, Appl	C 386	12.6	66.3	1188	6	US-10-953-349-27644	Sequence 27644, A
C 314	12.8	67.4	9737	1	US-09-484-331-23	Sequence 23, Appl	C 387	12.6	66.3	1188	6	US-11-217-529-5844	Sequence 5844, Ap
C 315	12.8	67.4	9737	1	US-09-484-331-28	Sequence 28, Appl	C 388	12.6	66.3	1197	6	US-10-953-349-14514	Sequence 14514, A
C 316	12.8	67.4	9871	1	US-09-484-331-24	Sequence 24, Appl	C 389	12.6	66.3	1209	6	US-10-471-571A-3933	Sequence 3933, Ap
C 317	12.8	67.4	10060	1	US-09-484-331-25	Sequence 25, Appl	C 390	12.6	66.3	1214	6	US-10-953-349-19083	Sequence 19083, A

C 391	12.6	66.3	1215	7	US-11-217-529-79049	Sequence 79049, A	C 464	12.6	66.3	2613	7	US-11-217-529-2125	Sequence 2125, Ap
C 392	12.6	66.3	1220	6	US-10-953-349-19685	Sequence 19685, A	C 465	12.6	66.3	2694	7	US-11-217-529-2262	Sequence 2262, Ap
C 393	12.6	66.3	1224	6	US-10-953-349-24547	Sequence 24547, A	C 466	12.6	66.3	2700	7	US-11-217-529-79652	Sequence 79652, A
C 394	12.6	66.3	1233	7	US-11-217-529-75482	Sequence 75482, A	C 467	12.6	66.3	2845	7	US-11-293-697-1481	Sequence 1481, Ap
C 395	12.6	66.3	1251	7	US-11-217-529-1972	Sequence 1972, Ap	C 468	12.6	66.3	2964	7	US-11-217-529-76841	Sequence 76841, A
C 396	12.6	66.3	1257	7	US-11-217-529-79920	Sequence 79920, Ap	C 469	12.6	66.3	2973	7	US-11-217-529-1745	Sequence 1745, Ap
C 397	12.6	66.3	1263	6	US-10-953-349-23655	Sequence 23655, A	C 470	12.6	66.3	3012	7	US-11-217-529-79922	Sequence 79922, Ap
C 398	12.6	66.3	1305	7	US-11-217-529-2579	Sequence 2579, Ap	C 471	12.6	66.3	3019	7	US-11-293-697-1626	Sequence 1626, Ap
C 399	12.6	66.3	1317	6	US-10-953-349-3399	Sequence 3399, Ap	C 472	12.6	66.3	3059	7	US-11-293-697-1913	Sequence 1913, Ap
C 400	12.6	66.3	1328	6	US-10-953-349-5258	Sequence 5258, Ap	C 473	12.6	66.3	3060	6	US-10-953-349-10946	Sequence 10946, A
C 401	12.6	66.3	1330	6	US-10-953-349-14318	Sequence 14318, A	C 474	12.6	66.3	3071	7	US-11-293-697-1990	Sequence 1990, Ap
C 402	12.6	66.3	1334	6	US-10-953-349-9228	Sequence 9228, Ap	C 475	12.6	66.3	3165	7	US-11-217-529-5854	Sequence 5854, Ap
C 403	12.6	66.3	1357	6	US-10-953-349-5349	Sequence 5349, Ap	C 476	12.6	66.3	3246	7	US-11-293-697-1216	Sequence 1216, Ap
C 404	12.6	66.3	1362	7	US-11-217-529-174096	Sequence 174096, A	C 477	12.6	66.3	3246	7	US-11-285-701-3	Sequence 3, Appli
C 405	12.6	66.3	1380	6	US-10-471-571A-4403	Sequence 4403, Ap	C 478	12.6	66.3	3293	7	US-11-285-701-2	Sequence 2, Appli
C 406	12.6	66.3	1395	6	US-10-953-349-3850	Sequence 3850, Ap	C 479	12.6	66.3	3323	6	US-10-953-349-1347	Sequence 1347, Ap
C 407	12.6	66.3	1397	6	US-10-953-349-2657	Sequence 2657, Ap	C 480	12.6	66.3	3357	7	US-11-217-529-77053	Sequence 77053, A
C 408	12.6	66.3	1412	6	US-10-953-349-10750	Sequence 10750, A	C 481	12.6	66.3	3435	6	US-10-953-349-4556	Sequence 4556, Ap
C 409	12.6	66.3	1413	7	US-11-217-529-2438	Sequence 2438, Ap	C 482	12.6	66.3	3567	7	US-11-311-778-19	Sequence 19, Appl
C 410	12.6	66.3	1414	6	US-10-953-349-5479	Sequence 5479, Ap	C 483	12.6	66.3	3651	6	US-10-471-571A-4941	Sequence 4941, Ap
C 411	12.6	66.3	1425	6	US-10-953-349-22922	Sequence 22922, A	C 484	12.6	66.3	3663	7	US-11-217-529-1369	Sequence 1369, Ap
C 412	12.6	66.3	1431	7	US-11-217-529-4228	Sequence 4228, Ap	C 485	12.6	66.3	3791	6	US-10-505-928-214	Sequence 214, App
C 413	12.6	66.3	1434	7	US-11-217-529-5339	Sequence 5339, Ap	C 486	12.6	66.3	3921	7	US-11-293-697-686	Sequence 686, App
C 414	12.6	66.3	1452	7	US-11-217-529-2756	Sequence 2756, Ap	C 487	12.6	66.3	3930	7	US-11-217-529-76026	Sequence 76026, A
C 415	12.6	66.3	1459	7	US-11-331-940-2	Sequence 2, Appli	C 488	12.6	66.3	4277	6	US-10-953-349-4544	Sequence 4544, Ap
C 416	12.6	66.3	1461	7	US-11-217-529-2101	Sequence 2101, Ap	C 489	12.6	66.3	5040	7	US-11-217-529-77889	Sequence 77889, A
C 417	12.6	66.3	1524	7	US-11-217-529-479	Sequence 479, App	C 490	12.6	66.3	5358	6	US-11-217-529-78153	Sequence 78153, A
C 418	12.6	66.3	1545	6	US-10-953-349-18450	Sequence 18450, A	C 491	12.6	66.3	5524	6	US-10-473-173-51	Sequence 51, Appl
C 419	12.6	66.3	1557	6	US-10-953-349-8484	Sequence 8484, Ap	C 492	12.6	66.3	5637	7	US-11-217-529-1074	Sequence 1074, Ap
C 420	12.6	66.3	1560	6	US-10-953-349-11211	Sequence 11211, A	C 493	12.6	66.3	5925	6	US-10-511-937-591	Sequence 591, App
C 421	12.6	66.3	1563	6	US-10-953-349-22887	Sequence 22887, A	C 494	12.6	66.3	7714	1	US-09-484-331-26	Sequence 26, Appl
C 422	12.6	66.3	1563	7	US-11-217-529-1045	Sequence 1045, Ap	C 495	12.6	66.3	9321	7	US-11-217-529-186	Sequence 186, App
C 423	12.6	66.3	1566	7	US-11-217-529-79698	Sequence 79698, A	C 496	12.6	66.3	16351	6	US-10-501-834-217	Sequence 217, App
C 424	12.6	66.3	1620	6	US-10-953-349-22250	Sequence 22250, A	C 497	12.4	65.3	25	7	US-11-217-529-55728	Sequence 55728, A
C 425	12.6	66.3	1638	7	US-11-217-529-2526	Sequence 2526, Ap	C 498	12.4	65.3	25	7	US-11-217-529-55732	Sequence 55732, A
C 426	12.6	66.3	1641	6	US-10-953-349-3315	Sequence 3315, Ap	C 499	12.4	65.3	25	7	US-11-217-529-98380	Sequence 98380, A
C 427	12.6	66.3	1666	6	US-10-953-349-37524	Sequence 37524, A	C 500	12.4	65.3	25	7	US-11-217-529-105242	Sequence 105242, A
C 428	12.6	66.3	1667	6	US-10-953-349-22972	Sequence 22972, A	C 501	12.4	65.3	69	7	US-11-217-529-4480	Sequence 4480, Ap
C 429	12.6	66.3	1673	6	US-11-217-529-76313	Sequence 76313, A	C 502	12.4	65.3	249	7	US-11-217-529-76741	Sequence 76741, A
C 430	12.6	66.3	1713	6	US-10-953-349-20667	Sequence 20667, A	C 503	12.4	65.3	372	7	US-11-217-529-6741	Sequence 2590, Ap
C 431	12.6	66.3	1737	6	US-10-953-349-14590	Sequence 14590, A	C 504	12.4	65.3	419	6	US-10-488-619-2590	Sequence 2590, Ap
C 432	12.6	66.3	1805	6	US-10-953-349-38219	Sequence 38219, A	C 505	12.4	65.3	486	6	US-11-217-529-77354	Sequence 77354, A
C 433	12.6	66.3	1817	6	US-10-953-349-34575	Sequence 34575, A	C 506	12.4	65.3	498	6	US-10-953-349-6839	Sequence 6839, Ap
C 434	12.6	66.3	1852	6	US-10-953-349-10311	Sequence 10311, A	C 507	12.4	65.3	551	6	US-10-488-619-2591	Sequence 2591, Ap
C 435	12.6	66.3	1889	6	US-10-953-349-23864	Sequence 23864, A	C 508	12.4	65.3	609	7	US-11-217-529-81492	Sequence 81492, A
C 436	12.6	66.3	1944	7	US-11-217-529-75861	Sequence 75861, A	C 509	12.4	65.3	612	7	US-11-217-529-77395	Sequence 77395, A
C 437	12.6	66.3	1948	6	US-10-953-349-10317	Sequence 10317, A	C 510	12.4	65.3	674	6	US-10-953-349-6905	Sequence 6905, Ap
C 438	12.6	66.3	1977	7	US-11-217-529-3395	Sequence 3395, Ap	C 511	12.4	65.3	876	6	US-10-953-349-489	Sequence 489, App
C 439	12.6	66.3	1980	7	US-11-217-529-3670	Sequence 3670, Ap	C 512	12.4	65.3	876	7	US-11-217-529-80000	Sequence 80000, A
C 440	12.6	66.3	2022	7	US-11-217-529-3785	Sequence 3785, Ap	C 513	12.4	65.3	936	7	US-11-217-529-75485	Sequence 75485, A
C 441	12.6	66.3	2022	7	US-11-267-837A-5	Sequence 4, Appli	C 514	12.4	65.3	1005	7	US-11-217-529-140	Sequence 140, App
C 442	12.6	66.3	2022	7	US-11-267-837A-4	Sequence 5, Appli	C 515	12.4	65.3	1029	6	US-10-953-349-32111	Sequence 32111, A
C 443	12.6	66.3	2067	6	US-10-953-349-3371	Sequence 3371, Ap	C 516	12.4	65.3	1041	7	US-11-217-529-79470	Sequence 79470, A
C 444	12.6	66.3	2075	6	US-10-953-349-8725	Sequence 8725, Ap	C 517	12.4	65.3	1104	7	US-11-217-529-2348	Sequence 2348, Ap
C 445	12.6	66.3	2079	6	US-10-953-349-5341	Sequence 5341, Ap	C 518	12.4	65.3	1104	6	US-11-217-529-78541	Sequence 78541, A
C 446	12.6	66.3	2079	7	US-11-217-529-2471	Sequence 2471, Ap	C 519	12.4	65.3	1203	6	US-10-471-571A-2671	Sequence 2671, Ap
C 447	12.6	66.3	2145	7	US-11-293-697-9	Sequence 9, Appli	C 520	12.4	65.3	1213	6	US-10-953-349-257	Sequence 257, App
C 448	12.6	66.3	2151	7	US-11-217-529-2775	Sequence 2775, Ap	C 521	12.4	65.3	1290	6	US-10-953-349-110853	Sequence 10853, A
C 449	12.6	66.3	2276	6	US-10-953-349-3916	Sequence 3916, Ap	C 522	12.4	65.3	1295	6	US-10-953-349-11518	Sequence 11518, A
C 450	12.6	66.3	2280	7	US-11-217-529-75361	Sequence 75361, A	C 523	12.4	65.3	1308	7	US-11-217-529-1529	Sequence 1529, Ap
C 451	12.6	66.3	2284	7	US-11-293-697-934	Sequence 934, App	C 524	12.4	65.3	1332	6	US-10-953-349-13736	Sequence 13736, A
C 452	12.6	66.3	2308	7	US-11-293-697-1274	Sequence 1274, Ap	C 525	12.4	65.3	1347	6	US-10-953-349-23316	Sequence 23316, A
C 453	12.6	66.3	2346	7	US-11-217-529-2131	Sequence 2131, Ap	C 526	12.4	65.3	1402	6	US-10-953-349-5210	Sequence 5210, Ap
C 454	12.6	66.3	2359	6	US-10-953-349-8431	Sequence 8431, Ap	C 527	12.4	65.3	1433	6	US-10-953-349-21194	Sequence 21194, A
C 455	12.6	66.3	2374	7	US-11-293-697-112	Sequence 112, App	C 528	12.4	65.3	1436	6	US-10-953-349-19256	Sequence 19256, A
C 456	12.6	66.3	2394	7	US-11-217-529-6126	Sequence 6126, Ap	C 529	12.4	65.3	1500	7	US-10-953-349-5405	Sequence 5405, Ap
C 457	12.6	66.3	2400	6	US-10-511-937-2827	Sequence 2827, Ap	C 530	12.4	65.3	1524	7	US-11-293-697-978	Sequence 978, App
C 458	12.6	66.3	2458	7	US-11-293-697-347	Sequence 347, App	C 531	12.4	65.3	1542	7	US-11-217-529-1926	Sequence 1926, Ap
C 459	12.6	66.3	2482	7	US-11-293-697-1488	Sequence 1488, Ap	C 532	12.4	65.3	1542	7	US-11-217-529-82420	Sequence 82420, A
C 460	12.6	66.3	2516	7	US-11-293-697-45	Sequence 45, Appl	C 533	12.4	65.3	1672	7	US-11-101-316-17	Sequence 17, Appl
C 461	12.6	66.3	2517	6	US-10-953-349-11366	Sequence 11366, A	C 534	12.4	65.3	1677	6	US-10-953-349-37561	Sequence 37561, A
C 462	12.6	66.3	2545	6	US-10-511-937-440	Sequence 440, App	C 535	12.4	65.3	1716	7	US-11-217-529-82504	Sequence 82504, A
C 463	12.6	66.3	2545	7	US-11-301-554-1900	Sequence 1900, Ap	C 536	12.4	65.3	1918	6	US-10-953-349-24339	Sequence 24339, A

537	12.4	65.3	1987	7	US-11-293-697-2224	Sequence 2224, Ap	c 610	12.2	64.2	762	7	US-11-217-529-1258	Sequence 1258, Ap
538	12.4	65.3	2067	6	US-10-953-349-5770	Sequence 5770, Ap	c 611	12.2	64.2	774	6	US-10-471-571A-3525	Sequence 3525, Ap
c 539	12.4	65.3	2163	7	US-11-217-529-4479	Sequence 4479, Ap	c 612	12.2	64.2	783	6	US-10-953-349-5500	Sequence 5500, Ap
c 540	12.4	65.3	2249	7	US-11-293-697-722	Sequence 722, Ap	c 613	12.2	64.2	792	7	US-11-217-529-7801	Sequence 7801, A
c 541	12.4	65.3	2265	7	US-11-217-529-78929	Sequence 78929, Ap	c 614	12.2	64.2	803	6	US-10-488-619-1428	Sequence 1428, Ap
c 542	12.4	65.3	2295	7	US-11-293-697-946	Sequence 946, Ap	c 615	12.2	64.2	804	6	US-10-953-349-35931	Sequence 35931, A
c 543	12.4	65.3	3126	7	US-11-251-340-4	Sequence 4, Appl	c 616	12.2	64.2	819	6	US-10-471-571A-3739	Sequence 3739, Ap
c 544	12.4	65.3	3296	7	US-11-293-697-688	Sequence 688, Ap	c 617	12.2	64.2	834	7	US-11-217-529-4582	Sequence 4582, Ap
c 545	12.4	65.3	3495	6	US-10-953-349-39546	Sequence 39546, A	c 618	12.2	64.2	844	6	US-10-953-349-38659	Sequence 38659, A
c 546	12.4	65.3	3667	7	US-11-293-697-33	Sequence 33, Appl	c 619	12.2	64.2	851	6	US-10-953-349-6165	Sequence 6165, Ap
c 547	12.4	65.3	4674	7	US-11-217-529-1182	Sequence 1182, Ap	c 620	12.2	64.2	852	7	US-11-217-529-78919	Sequence 78919, A
548	12.2	64.2	25	7	US-11-217-529-24087	Sequence 24087, A	c 621	12.2	64.2	855	6	US-11-217-529-5066	Sequence 5066, Ap
549	12.2	64.2	25	7	US-11-217-529-30382	Sequence 30382, A	c 622	12.2	64.2	856	6	US-10-953-349-17527	Sequence 17527, A
550	12.2	64.2	25	7	US-11-217-529-94099	Sequence 94099, A	c 623	12.2	64.2	857	6	US-10-953-349-32928	Sequence 32928, A
551	12.2	64.2	25	7	US-11-217-529-107138	Sequence 107138, A	c 624	12.2	64.2	858	7	US-11-217-529-173833	Sequence 173833
c 552	12.2	64.2	25	7	US-11-217-529-174769	Sequence 174769, A	c 625	12.2	64.2	867	7	US-11-217-529-2211	Sequence 2211, Ap
c 553	12.2	64.2	25	7	US-11-217-529-174771	Sequence 174771, A	c 626	12.2	64.2	883	6	US-10-953-349-10460	Sequence 10460, A
554	12.2	64.2	195	6	US-10-554-711-195	Sequence 195, App	c 627	12.2	64.2	885	7	US-11-217-529-77274	Sequence 77274, A
c 555	12.2	64.2	51	7	US-11-245-248-281	Sequence 281, App	c 628	12.2	64.2	888	6	US-10-488-619-2356	Sequence 2356, Ap
c 556	12.2	64.2	51	7	US-11-245-248-335	Sequence 335, App	c 629	12.2	64.2	891	7	US-11-217-529-4427	Sequence 4427, Ap
c 557	12.2	64.2	84	7	US-11-217-529-173140	Sequence 173140, A	c 630	12.2	64.2	897	6	US-10-471-571A-3807	Sequence 3807, Ap
c 558	12.2	64.2	110	7	US-11-194-055-241	Sequence 241, App	c 631	12.2	64.2	897	7	US-11-217-529-5409	Sequence 5409, Ap
c 559	12.2	64.2	126	6	US-10-471-571A-3727	Sequence 3727, Ap	c 632	12.2	64.2	901	6	US-10-953-349-18012	Sequence 18012, A
c 560	12.2	64.2	186	7	US-11-217-529-191000	Sequence 191000, A	c 633	12.2	64.2	906	6	US-10-471-571A-3187	Sequence 3187, Ap
c 561	12.2	64.2	216	6	US-10-471-571A-3415	Sequence 3415, Ap	c 634	12.2	64.2	914	6	US-10-953-349-15932	Sequence 15932, A
c 562	12.2	64.2	237	6	US-10-471-571A-4727	Sequence 4727, Ap	c 635	12.2	64.2	921	7	US-11-217-529-76096	Sequence 76096, A
c 563	12.2	64.2	294	6	US-10-471-571A-2869	Sequence 2869, Ap	c 636	12.2	64.2	925	6	US-10-953-349-19602	Sequence 19602, A
c 564	12.2	64.2	309	7	US-11-301-554-1222	Sequence 1222, Ap	c 637	12.2	64.2	933	6	US-10-953-349-37077	Sequence 37077, A
c 565	12.2	64.2	312	7	US-11-217-529-166750	Sequence 166750, A	c 638	12.2	64.2	942	7	US-11-217-529-82682	Sequence 82682, A
c 566	12.2	64.2	321	7	US-11-301-554-476	Sequence 476, App	c 639	12.2	64.2	951	6	US-10-471-571A-973	Sequence 973, App
c 567	12.2	64.2	324	7	US-11-217-529-173586	Sequence 173586, A	c 640	12.2	64.2	960	6	US-10-953-349-13653	Sequence 13653, A
c 568	12.2	64.2	333	7	US-11-217-529-1041	Sequence 1041, Ap	c 641	12.2	64.2	963	7	US-11-217-529-80262	Sequence 80262, A
c 569	12.2	64.2	345	7	US-11-217-529-191075	Sequence 191075, A	c 642	12.2	64.2	972	6	US-10-471-571A-2991	Sequence 2991, Ap
c 570	12.2	64.2	347	6	US-10-488-619-634	Sequence 634, App	c 643	12.2	64.2	975	7	US-11-217-529-80594	Sequence 80594, A
c 571	12.2	64.2	347	7	US-11-222-810-1	Sequence 1, Appl	c 644	12.2	64.2	978	7	US-11-217-529-258	Sequence 258, App
c 572	12.2	64.2	347	7	US-11-222-810-3	Sequence 3, Appl	c 645	12.2	64.2	978	7	US-11-217-529-75365	Sequence 75365, A
c 573	12.2	64.2	363	7	US-11-217-529-4078	Sequence 4078, Ap	c 646	12.2	64.2	978	7	US-11-217-529-81375	Sequence 81375, A
c 574	12.2	64.2	375	7	US-11-217-529-81887	Sequence 81887, A	c 647	12.2	64.2	984	6	US-10-471-571A-1451	Sequence 1451, Ap
c 575	12.2	64.2	381	6	US-10-471-571A-2657	Sequence 2657, Ap	c 648	12.2	64.2	990	7	US-11-217-529-6230	Sequence 6230, Ap
c 576	12.2	64.2	390	7	US-11-217-529-173480	Sequence 173480, A	c 649	12.2	64.2	990	7	US-11-217-529-77231	Sequence 77231, A
c 577	12.2	64.2	402	7	US-11-217-529-77048	Sequence 77048, A	c 650	12.2	64.2	993	6	US-10-471-571A-2969	Sequence 2969, Ap
c 578	12.2	64.2	435	7	US-11-217-529-174533	Sequence 174533, A	c 651	12.2	64.2	1005	7	US-11-217-529-4665	Sequence 4665, Ap
c 579	12.2	64.2	438	6	US-10-471-571A-3115	Sequence 3115, Ap	c 652	12.2	64.2	1014	7	US-11-217-529-1646	Sequence 1646, Ap
c 580	12.2	64.2	444	6	US-10-953-349-15942	Sequence 15942, A	c 653	12.2	64.2	1038	7	US-11-217-529-77823	Sequence 77823, A
c 581	12.2	64.2	446	6	US-10-953-349-15942	Sequence 15942, A	c 654	12.2	64.2	1042	6	US-10-953-349-39084	Sequence 39084, A
c 582	12.2	64.2	450	7	US-11-217-529-5154	Sequence 5154, Ap	c 655	12.2	64.2	1044	6	US-11-217-529-79320	Sequence 79320, A
c 583	12.2	64.2	460	7	US-11-301-554-709	Sequence 709, App	c 656	12.2	64.2	1059	6	US-10-953-349-32351	Sequence 32351, A
c 584	12.2	64.2	468	7	US-11-301-554-539	Sequence 539, App	c 657	12.2	64.2	1062	7	US-11-217-529-5471	Sequence 5471, Ap
c 585	12.2	64.2	501	7	US-11-183-218-19	Sequence 19, Appl	c 658	12.2	64.2	1062	7	US-11-217-529-81237	Sequence 81237, Ap
c 586	12.2	64.2	506	6	US-10-488-619-2998	Sequence 2998, Ap	c 659	12.2	64.2	1075	6	US-10-953-349-11726	Sequence 11726, A
c 587	12.2	64.2	506	6	US-10-471-571A-637	Sequence 637, App	c 660	12.2	64.2	1098	7	US-11-217-529-5465	Sequence 5465, Ap
c 588	12.2	64.2	558	6	US-10-488-619-2802	Sequence 2802, Ap	c 661	12.2	64.2	1098	7	US-11-217-529-77660	Sequence 77660, A
c 589	12.2	64.2	561	7	US-11-301-554-1341	Sequence 1341, Ap	c 662	12.2	64.2	1110	6	US-10-471-571A-5027	Sequence 5027, Ap
c 590	12.2	64.2	570	6	US-10-471-571A-4957	Sequence 4957, Ap	c 663	12.2	64.2	1110	6	US-11-217-529-4050	Sequence 4050, Ap
c 591	12.2	64.2	578	6	US-10-488-619-1135	Sequence 1135, Ap	c 664	12.2	64.2	1125	6	US-11-217-529-76126	Sequence 76126, A
c 592	12.2	64.2	584	7	US-11-217-529-77047	Sequence 77047, A	c 665	12.2	64.2	1129	6	US-10-953-349-22496	Sequence 22496, A
c 593	12.2	64.2	594	7	US-11-217-529-81783	Sequence 81783, A	c 666	12.2	64.2	1143	7	US-11-217-529-173391	Sequence 173391, A
c 594	12.2	64.2	606	7	US-11-217-529-5255	Sequence 5255, Ap	c 667	12.2	64.2	1146	7	US-11-217-529-190964	Sequence 190964, A
c 595	12.2	64.2	618	7	US-11-217-529-76793	Sequence 76793, A	c 668	12.2	64.2	1149	7	US-11-217-529-78468	Sequence 78468, A
c 596	12.2	64.2	624	6	US-10-471-571A-439	Sequence 439, App	c 669	12.2	64.2	1152	7	US-11-217-529-190877	Sequence 190877, A
c 597	12.2	64.2	632	6	US-10-953-349-37785	Sequence 37785, A	c 670	12.2	64.2	1156	6	US-10-953-349-27995	Sequence 27995, A
c 598	12.2	64.2	634	6	US-10-488-619-2166	Sequence 2166, Ap	c 671	12.2	64.2	1164	7	US-11-217-529-497	Sequence 497, App
c 599	12.2	64.2	638	6	US-10-537-455-3	Sequence 3, Appl	c 672	12.2	64.2	1165	6	US-10-953-349-23368	Sequence 23368, A
c 600	12.2	64.2	643	6	US-10-488-619-2957	Sequence 2957, Ap	c 673	12.2	64.2	1165	6	US-10-953-349-10445	Sequence 10445, A
c 601	12.2	64.2	659	6	US-10-953-349-16589	Sequence 16589, A	c 674	12.2	64.2	1169	6	US-10-953-349-38273	Sequence 38273, A
c 602	12.2	64.2	682	6	US-10-953-349-13084	Sequence 13084, A	c 675	12.2	64.2	1179	6	US-10-953-349-18399	Sequence 18399, A
c 603	12.2	64.2	687	7	US-11-217-529-79096	Sequence 79096, A	c 676	12.2	64.2	1182	7	US-11-217-529-2391	Sequence 2391, Ap
c 604	12.2	64.2	693	6	US-10-488-619-2853	Sequence 2853, Ap	c 677	12.2	64.2	1185	7	US-11-217-529-76810	Sequence 76810, A
c 605	12.2	64.2	708	7	US-11-217-529-174078	Sequence 174078, A	c 678	12.2	64.2	1193	6	US-10-511-937-445	Sequence 445, App
c 606	12.2	64.2	714	7	US-11-217-529-76479	Sequence 76479, Ap	c 679	12.2	64.2	1194	6	US-10-471-571A-661	Sequence 661, App
c 607	12.2	64.2	716	6	US-10-953-349-1623	Sequence 1623, Ap	c 680	12.2	64.2	1202	6	US-10-953-349-38715	Sequence 38715, A
c 608	12.2	64.2	717	7	US-11-217-529-1101	Sequence 1101, Ap	c 681	12.2	64.2	1203	7	US-11-217-529-78331	Sequence 78331, A
c 609	12.2	64.2	735	6	US-10-953-349-8502	Sequence 8502, Ap	c 682	12.2	64.2	1208	6	US-10-953-349-19298	Sequence 19298, A

C 829	12.2	64.2	2375	6	US-10-953-349-11035	Sequence 11035, A	C 902	12.2	64.2	5957	6	US-10-473-173-106	Sequence 106, App
C 830	12.2	64.2	2387	7	US-11-217-529-11035	Sequence 50, Appl	C 903	12.2	64.2	6045	7	US-11-217-529-1312	Sequence 1312, Ap
C 831	12.2	64.2	2436	7	US-11-217-529-1652	Sequence 1522, A	C 904	12.2	64.2	6317	7	US-10-511-937-387	Sequence 387, App
C 832	12.2	64.2	2436	7	US-11-217-529-82283	Sequence 8228, Ap	C 905	12.2	64.2	6891	7	US-11-251-724-2	Sequence 2, Appl
C 833	12.2	64.2	2440	7	US-11-293-697-1135	Sequence 1135, Ap	C 906	12.2	64.2	7131	7	US-11-217-529-77074	Sequence 77074, A
C 834	12.2	64.2	2479	7	US-11-293-697-1966	Sequence 1966, Ap	C 907	12.2	64.2	7254	7	US-11-217-529-1173	Sequence 1173, Ap
C 835	12.2	64.2	2498	7	US-11-293-697-1092	Sequence 1092, Ap	C 908	12.2	64.2	7873	6	US-10-485-397-7	Sequence 7, Appl
C 836	12.2	64.2	2537	6	US-10-505-928-149	Sequence 149, App	C 909	12.2	64.2	7943	6	US-10-485-397-8	Sequence 8, Appl
C 837	12.2	64.2	2559	7	US-11-293-697-83	Sequence 83, Appl	C 910	12.2	64.2	8090	7	US-11-226-605-86	Sequence 86, Appl
C 838	12.2	64.2	2562	7	US-11-217-529-4523	Sequence 4523, Ap	C 911	12.2	64.2	8448	7	US-11-145-3078-166	Sequence 166, App
C 839	12.2	64.2	2568	7	US-11-217-529-77262	Sequence 77262, A	C 912	12.2	64.2	9353	6	US-10-480-962-38	Sequence 38, Appl
C 840	12.2	64.2	2572	1	US-09-949-925-36	Sequence 36, Appl	C 913	12.2	64.2	9449	6	US-10-480-962-39	Sequence 39, Appl
C 841	12.2	64.2	2607	7	US-11-217-529-3100	Sequence 3100, Ap	C 914	12.2	64.2	12342	7	US-11-217-529-4644	Sequence 4644, Ap
C 842	12.2	64.2	2616	7	US-11-217-529-3604	Sequence 3604, Ap	C 915	12.2	64.2	15304	6	US-10-505-928-828	Sequence 828, App
C 843	12.2	64.2	2638	7	US-11-293-697-879	Sequence 879, Ap	C 916	12.2	64.2	37426	6	US-10-473-173-32	Sequence 32, Appl
C 844	12.2	64.2	2645	7	US-11-181-115-25	Sequence 25, Appl	C 917	12.2	64.2	56580	6	US-10-553-298-1	Sequence 1, Appl
C 845	12.2	64.2	2658	7	US-11-293-697-807	Sequence 807, App	C 918	12.2	64.2	70665	6	US-10-505-928-596	Sequence 596, App
C 846	12.2	64.2	2650	7	US-11-293-697-241	Sequence 241, App	C 919	12.2	64.2	70665	6	US-10-505-928-596	Sequence 596, App
C 847	12.2	64.2	2655	6	US-10-473-173-96	Sequence 96, Appl	C 920	12.2	64.2	118899	7	US-11-189-279-64	Sequence 64, Appl
C 848	12.2	64.2	2663	6	US-10-953-349-8793	Sequence 8793, Ap	C 921	12.2	64.2	118899	7	US-10-505-928-607	Sequence 607, App
C 849	12.2	64.2	2682	7	US-11-217-529-571	Sequence 571, App	C 922	12.2	64.2	135090	6	US-10-505-928-607	Sequence 607, App
C 850	12.2	64.2	2694	7	US-11-217-529-5531	Sequence 5531, Ap	C 923	12.2	64.2	394191	6	US-10-506-549-3	Sequence 3, Appl
C 851	12.2	64.2	2733	7	US-11-217-529-76516	Sequence 76516, A	C 924	12.2	63.2	25	7	US-11-217-529-137480	Sequence 137480,
C 852	12.2	64.2	2760	7	US-11-217-529-80783	Sequence 80783, A	C 925	12.2	63.2	25	7	US-11-217-529-142134	Sequence 142134,
C 853	12.2	64.2	2778	7	US-11-217-529-1916	Sequence 1916, Ap	C 926	12.2	63.2	631	6	US-11-301-554-566	Sequence 566, App
C 854	12.2	64.2	2808	7	US-11-181-115-23	Sequence 23, Appl	C 927	12.2	63.2	751	6	US-10-953-349-31486	Sequence 31486, A
C 855	12.2	64.2	2808	7	US-11-181-115-27	Sequence 27, Appl	C 928	12.2	63.2	1172	6	US-10-953-349-6628	Sequence 6628, Ap
C 856	12.2	64.2	2808	7	US-11-181-115-41	Sequence 41, Appl	C 929	12.2	63.2	1177	6	US-10-953-349-8634	Sequence 8634, Ap
C 857	12.2	64.2	2811	7	US-11-217-529-3053	Sequence 3053, Ap	C 930	12.2	63.2	1204	6	US-10-953-349-29986	Sequence 29986, A
C 858	12.2	64.2	2886	7	US-11-217-529-75510	Sequence 75510, A	C 931	12.2	63.2	1254	7	US-11-217-529-81803	Sequence 81803, A
C 859	12.2	64.2	2901	7	US-11-217-529-2168	Sequence 2168, Ap	C 932	12.2	63.2	1284	7	US-11-217-529-1514	Sequence 1514, A
C 860	12.2	64.2	2916	7	US-11-217-529-81609	Sequence 81609, A	C 933	12.2	63.2	1284	7	US-11-217-529-77448	Sequence 77448, A
C 861	12.2	64.2	2976	6	US-10-953-349-37847	Sequence 37847, A	C 934	12.2	63.2	1454	6	US-10-953-349-2764	Sequence 2764, Ap
C 862	12.2	64.2	2980	7	US-11-293-697-465	Sequence 465, App	C 935	12.2	63.2	1515	6	US-11-217-529-3477	Sequence 3477, Ap
C 863	12.2	64.2	2983	7	US-11-293-697-589	Sequence 589, App	C 936	12.2	63.2	1620	6	US-10-471-571A-4207	Sequence 4207, Ap
C 864	12.2	64.2	2988	7	US-11-217-529-80377	Sequence 80377, A	C 937	12.2	63.2	1631	6	US-10-953-349-12559	Sequence 12559, A
C 865	12.2	64.2	2997	7	US-11-217-529-1582	Sequence 1582, Ap	C 938	12.2	63.2	1725	7	US-11-217-529-81352	Sequence 81352, A
C 866	12.2	64.2	3012	7	US-11-217-529-3912	Sequence 3912, Ap	C 939	12.2	63.2	1737	7	US-11-217-529-1440	Sequence 1440, Ap
C 867	12.2	64.2	3012	7	US-11-293-697-720	Sequence 720, App	C 940	12.2	63.2	1937	1	US-09-949-925-61	Sequence 61, Appl
C 868	12.2	64.2	3045	7	US-11-293-697-1338	Sequence 1338, Ap	C 941	12.2	63.2	2049	7	US-11-217-529-79820	Sequence 79820, A
C 869	12.2	64.2	3069	7	US-11-217-529-80037	Sequence 80027, A	C 942	12.2	63.2	2076	7	US-11-217-529-352	Sequence 352, App
C 870	12.2	64.2	3128	7	US-11-293-697-1259	Sequence 1259, Ap	C 943	12.2	63.2	2076	7	US-11-106-014-9	Sequence 9, Appl
C 871	12.2	64.2	3198	6	US-10-953-349-8411	Sequence 8411, Ap	C 944	12.2	63.2	2082	7	US-11-217-529-2378	Sequence 2378, Ap
C 872	12.2	64.2	3252	7	US-11-318-813-3	Sequence 3, Appl	C 945	12.2	63.2	2139	6	US-10-511-937-338	Sequence 338, App
C 873	12.2	64.2	3271	7	US-11-222-810-22	Sequence 22, Appl	C 946	12.2	63.2	2168	7	US-11-293-697-2011	Sequence 2011, Ap
C 874	12.2	64.2	3271	7	US-11-293-697-244	Sequence 244, App	C 947	12.2	63.2	2847	7	US-11-293-697-403	Sequence 403, App
C 875	12.2	64.2	3278	7	US-11-145-3078-198	Sequence 198, App	C 948	12.2	63.2	3609	7	US-11-217-529-77868	Sequence 77868, A
C 876	12.2	64.2	3399	7	US-11-217-529-73	Sequence 73, Appl	C 949	12.2	63.2	4308	7	US-11-217-529-82409	Sequence 82409, A
C 877	12.2	64.2	3408	7	US-11-217-529-2270	Sequence 2270, Ap	C 950	12.2	63.2	4656	7	US-11-217-529-3130	Sequence 3130, Ap
C 878	12.2	64.2	3435	7	US-11-217-529-989	Sequence 989, App	C 951	12.2	63.2	5748	7	US-11-217-529-3601	Sequence 3601, Ap
C 879	12.2	64.2	3441	7	US-11-217-529-76893	Sequence 76893, A	C 952	12.2	63.2	5833	7	US-11-024-544A-139	Sequence 139, App
C 880	12.2	64.2	3485	7	US-11-145-3078-192	Sequence 192, App	C 953	12.2	63.2	5833	7	US-11-024-545-61	Sequence 61, Appl
C 881	12.2	64.2	3486	6	US-10-953-349-16524	Sequence 16524, A	C 954	12.2	63.2	5833	7	US-11-251-466-35	Sequence 35, Appl
C 882	12.2	64.2	3628	7	US-11-145-3078-15	Sequence 15, Appl	C 955	12.2	63.2	5833	7	US-11-254-173-46	Sequence 46, Appl
C 883	12.2	64.2	3855	7	US-11-217-529-79166	Sequence 79166, A	C 956	12.2	63.2	5833	7	US-11-264-784-131	Sequence 131, App
C 884	12.2	64.2	4015	7	US-11-293-697-627	Sequence 627, App	C 957	12.2	63.2	5833	7	US-11-264-737-159	Sequence 159, App
C 885	12.2	64.2	4021	7	US-11-293-697-472	Sequence 472, App	C 958	12.2	63.2	5833	7	US-11-265-761-133	Sequence 133, App
C 886	12.2	64.2	4098	6	US-10-510-508-2	Sequence 2, Appl	C 959	12.2	63.2	6457	7	US-11-264-784-117	Sequence 117, App
C 887	12.2	64.2	4110	6	US-10-471-571A-739	Sequence 739, App	C 960	12.2	63.2	6457	7	US-11-264-737-166	Sequence 166, App
C 888	12.2	64.2	4179	7	US-11-217-529-176	Sequence 176, App	C 961	12.2	63.2	6457	7	US-11-265-761-142	Sequence 142, App
C 889	12.2	64.2	4257	7	US-11-217-529-2743	Sequence 2743, Ap	C 962	12.2	63.2	6498	7	US-11-264-784-148	Sequence 148, App
C 890	12.2	64.2	4419	7	US-11-217-529-2484	Sequence 2484, Ap	C 963	12.2	63.2	6498	7	US-11-264-737-200	Sequence 200, App
C 891	12.2	64.2	4498	7	US-11-217-529-190990	Sequence 190990, A	C 964	12.2	63.2	6498	7	US-11-265-761-163	Sequence 163, App
C 892	12.2	64.2	4654	6	US-10-505-928-20	Sequence 20, Appl	C 965	12.2	63.2	6540	7	US-11-254-173-60	Sequence 60, Appl
C 893	12.2	64.2	4688	7	US-11-293-697-222	Sequence 222, App	C 966	12.2	63.2	6540	7	US-11-264-784-135	Sequence 135, App
C 894	12.2	64.2	4761	7	US-11-217-529-1821	Sequence 1821, Ap	C 967	12.2	63.2	6540	7	US-11-264-737-187	Sequence 187, App
C 895	12.2	64.2	4804	7	US-11-181-115-42	Sequence 42, Appl	C 968	12.2	63.2	6540	7	US-11-265-761-139	Sequence 139, App
C 896	12.2	64.2	4962	7	US-11-217-529-76536	Sequence 76536, A	C 969	12.2	63.2	6561	7	US-11-264-784-147	Sequence 147, App
C 897	12.2	64.2	5009	6	US-11-244-452-8	Sequence 8, Appl	C 970	12.2	63.2	6561	7	US-11-264-784-150	Sequence 150, App
C 898	12.2	64.2	5072	6	US-10-953-349-7998	Sequence 7998, Ap	C 971	12.2	63.2	6561	7	US-11-264-737-199	Sequence 199, App
C 899	12.2	64.2	5540	1	US-09-970-076-5	Sequence 5, Appl	C 972	12.2	63.2	6561	7	US-11-264-737-202	Sequence 202, App
C 900	12.2	64.2	5688	7	US-11-217-529-76708	Sequence 76708, A	C 973	12.2	63.2	6561	7	US-11-265-761-162	Sequence 162, App
C 901	12.2	64.2	5920	7	US-11-145-3078-185	Sequence 185, App	C 974	12.2	63.2	6561	7	US-11-265-761-165	Sequence 165, App

Sequence 57, Appl
Sequence 133, App
Sequence 185, App
Sequence 138, App
Sequence 168, App
Sequence 68, Appl
Sequence 53, Appl
Sequence 61, Appl
Sequence 140, Appl
Sequence 192, App
Sequence 155, App
Sequence 144, App
Sequence 153, App
Sequence 205, App
Sequence 169, App
Sequence 119, App
Sequence 167, App
Sequence 144, App
Sequence 154, App
Sequence 206, App
Sequence 169, App
Sequence 41, Appl
Sequence 138, App
Sequence 190, App
Sequence 153, App
Sequence 67, Appl

12 63.2 6912 7 US-11-254-173-57
12 63.2 6912 7 US-11-264-784-133
12 63.2 6912 7 US-11-264-737-185
12 63.2 6912 7 US-11-265-761-138
12 63.2 7323 7 US-11-024-544A-168
12 63.2 7323 7 US-11-024-545-68
12 63.2 7323 7 US-11-251-466-53
12 63.2 7323 7 US-11-254-173-61
12 63.2 7323 7 US-11-264-784-140
12 63.2 7323 7 US-11-264-737-192
12 63.2 7323 7 US-11-265-761-155
12 63.2 7822 7 US-11-024-544A-144
12 63.2 7822 7 US-11-264-784-153
12 63.2 7822 7 US-11-264-737-205
12 63.2 7822 7 US-11-265-761-168
12 63.2 7879 7 US-11-264-784-119
12 63.2 7879 7 US-11-264-737-167
12 63.2 7879 7 US-11-265-761-144
12 63.2 7930 7 US-11-264-784-154
12 63.2 7930 7 US-11-264-737-206
12 63.2 7930 7 US-11-265-761-169
12 63.2 8015 7 US-11-251-466-41
12 63.2 8015 7 US-11-264-784-138
12 63.2 8015 7 US-11-264-737-190
12 63.2 8015 7 US-11-265-761-153
1000 12 63.2 8084 7 US-11-024-545-67

ALIGNMENTS

RESULT 1
US-10-471-571A-1359/c
; Sequence 1359, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1359
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Staphylococcus aureus

Query Match 81.1%; Score 15.4; DB 6; Length 522;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCTTT 18
| | | | | | | | | |
Db 209 TTACTTCATAGCCTTT 193

RESULT 2
US-11-217-529-78062/c
; Sequence 78062, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285

; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78062
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78062

Query Match 81.1%; Score 15.4; DB 7; Length 2163;
Best Local Similarity 94.1%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
| | | | | | | | | |
Db 1662 TTCTTCATAGTCTTTG 1646

RESULT 3
US-11-313-450-17
; Sequence 17, Application US/11313450
; Publication No. US20060110778A1
; GENERAL INFORMATION:
; APPLICANT: Adorante, Joeseeph S.
; APPLICANT: Ehring, George R.
; TITLE OF INVENTION: High-Throughput Screens For Identifying
; TITLE OF INVENTION: Selective Persistent Sodium Channels Channel Blockers
; FILE REFERENCE: 17407 CIP (AP)
; CURRENT APPLICATION NUMBER: US/11/313,450
; CURRENT FILING DATE: 2005-12-19
; PRIOR APPLICATION NUMBER: 09/989797
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252771
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 5376
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match 81.1%; Score 15.4; DB 7; Length 5376;
Best Local Similarity 94.1%; Pred. No. 55;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCTTT 18
| | | | | | | | | |
Db 3857 TTACTTCGATGCTTT 3873

RESULT 4
US-10-953-349-5467/c
; Sequence 5467, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5467
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5467

Query Match 78.9%; Score 15; DB 6; Length 1446;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TACTTCATAGCTTTT 18
|||||
Db 1187 TACTTCATAGCTTTT 1173

RESULT 5

US-11-183-218-74/c
; Sequence 74, Application US/11183218
; Publication No. US2006008906A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DePree, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryne
; TITLE OF INVENTION: ERYTHROPOIETIN: REMODELING AND
; TITLE OF INVENTION: GLYCOCONJUGATION OF ERYTHROPOIETIN
; FILE REFERENCE: 040853-01-5083-US02
; CURRENT APPLICATION NUMBER: US/11/183.218
; CURRENT FILING DATE: 2005-07-15
; PRIOR FILING DATE: 2005-10-15
; PRIOR FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: PCT/US02/32263
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-28

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-183-218-74

Query Match 77.9%; Score 14.8; DB 7; Length 588;
Best Local Similarity 88.9%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTTTT 18
|||||
Db 565 CTTTACTTCTCAGCTTTT 548

RESULT 6

US-11-217-529-76788/c
; Sequence 76788, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO

; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217.529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76788
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76788

Query Match 77.9%; Score 14.8; DB 7; Length 654;
Best Local Similarity 88.9%; Pred. No. 82;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTTACTTCATAGCTTTTG 19
|||||
Db 578 TTGACTTCATAGCTCTG 561

RESULT 7

US-10-953-349-6204/c
; Sequence 6204, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6204
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-6204

Query Match 77.9%; Score 14.8; DB 6; Length 676;
Best Local Similarity 88.9%; Pred. No. 82;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTTACTTCATAGCTTTTG 19
|||||
Db 199 TTTTACTTCATAGCTCTG 182

RESULT 8

US-10-953-349-2872
; Sequence 2872, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2872
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2872

Query Match 77.9%; Score 14.8; DB 6; Length 956;

```
Best Local Similarity 88.9%; Pred. No. 86;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCCTTTG 19
   ||||| ||||| |||||
Db 440 TTTACTTCAAGTTTGTG 457

RESULT 9
US-11-217-529-78358/c
; Sequence 78358, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78358
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78358

Query Match 77.9%; Score 14.8; DB 7; Length 1212;
Best Local Similarity 88.9%; Pred. No. 88;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTT 18
   ||||| ||||| |||||
Db 1127 CTTTGCTTCATAGTGT 1110

RESULT 10
US-10-953-349-2801
; Sequence 2801, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2801
; LENGTH: 1432
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2801

Query Match 77.9%; Score 14.8; DB 6; Length 1432;
Best Local Similarity 88.9%; Pred. No. 90;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTT 18
   ||||| ||||| |||||
Db 57 CTTTCTTCATCGCTTT 74

RESULT 11
US-10-541-947-5/c
; Sequence 5, Application US/10541947
```

```
Publication No. US20060095980A1
; GENERAL INFORMATION:
; APPLICANT: North Carolina State University
; APPLICANT: Pettite, James
; APPLICANT: Pardue, Samuel
; TITLE OF INVENTION: DEPLETION OF ENDOGENOUS PRIMORDIAL GERM CELLS IN AVIAN SPECIES
; FILE REFERENCE: 297/204 PCT
; CURRENT APPLICATION NUMBER: US/10/541,947
; CURRENT FILING DATE: 2005-07-08
; PRIOR APPLICATION NUMBER: US 60/440,424
; PRIOR FILING DATE: 2003-01-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (174)..(1043)
US-10-541-947-5

Query Match 77.9%; Score 14.8; DB 6; Length 1882;
Best Local Similarity 88.9%; Pred. No. 92;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TTTACTTCATAGTCCTTTG 19
   ||||| ||||| |||||
Db 1065 TGTACTTCAAGTCTTTG 1048

RESULT 12
US-11-293-697-610
; Sequence 610, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 610
; LENGTH: 3551
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-610

Query Match 77.9%; Score 14.8; DB 7; Length 3551;
Best Local Similarity 88.9%; Pred. No. 99;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCCTTT 18
   ||||| ||||| |||||
Db 2987 CTTTAATTCATAGCCTTT 3004

RESULT 13
US-11-293-697-662
; Sequence 662, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
```



```
QY      2 TTTACTTCATAGTCTT 17
      ||| ||| ||| ||| |||
Db      314 TTTCTTCATAGTCTT 329

RESULT 18
US-10-953-349-11495/c
; Sequence 11495, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCED-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11495
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11495

Query Match      75.8%; Score 14.4; DB 6; Length 1457;
Best Local Similarity 93.8%; Pred. No. 1.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TTTACTTCATAGTCTT 17
      ||| ||| ||| ||| |||
Db      753 TTGACTTCATAGTCTT 738

RESULT 19
US-11-293-697-523/c
; Sequence 523, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 523
; LENGTH: 3689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-523

Query Match      75.8%; Score 14.4; DB 7; Length 3689;
Best Local Similarity 93.8%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 TACTTCATAGTCTTTG 19
      ||| ||| ||| ||| |||
Db      3428 TACTTCATAGTCTTTG 3413

RESULT 20
US-11-217-529-161319
; Sequence 161319, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
```

```
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 161319
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-161319

Query Match      74.7%; Score 14.2; DB 7; Length 25;
Best Local Similarity 84.2%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCTTTG 19
      ||| ||| ||| ||| |||
Db      3 CTTTGGTCATCGTCTTTG 21

RESULT 21
US-11-217-529-172215
; Sequence 172215, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 172215
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-172215

Query Match      74.7%; Score 14.2; DB 7; Length 25;
Best Local Similarity 84.2%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCTTTG 19
      ||| ||| ||| ||| |||
Db      2 CTTTCTTCGTAGTCTTTG 20

RESULT 22
US-11-217-529-166960
; Sequence 166960, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
```

```

; PRIOR APPLICATION NUMBER: US 10/932,182
;
; PRIOR FILING DATE: 2004-09-02
;
; NUMBER OF SEQ ID NOS: 197023
;
; SOFTWARE: PatentIn version 3.3
;
; SEQ ID NO 166960
;
; LENGTH: 152
;
; TYPE: DNA
;
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-166960

```

Query Match 74.7%; Score 14.2; DB 7; Length 192;
Best Local Similarity 84.2%; Pred. NO. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
||| ||| ||| ||| |||
Db 13 CTTTTCCTCGTAGTTTTTG 31

```

RESULT 23
US-11-217-529-80554/c
; Sequence 80554, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529

```

```

; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 80554
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-80554

```

Query Match 74.7%; Score 14.2; DB 7; Length 255;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTG 19
 ||| ||||| | |||||
Db 174 CTTGACTTCATCGTATTG 156

```

RESULT 24
US-10-953-349-10182/c
; Sequence 10182, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10182
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10182

```

Query Match 74.7%; Score 14.2; DB 6; Length 517;

```

Best Local Similarity 84.2%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGCTCTTTG 19
    ||||| || |||||
Db 85 CTTTCTCTCTTAGCTCTTTG 67

```

RESULT 25
 US-11-217-529-82353
 ; Sequence 82353, Application US/11217529
 ; Publication No. US20060099612A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNTORY LIMITED
 ; APPLICANT: NAKAO, YOSHIHIRO
 ; APPLICANT: NAKAMURA, NORIHIISA
 ; APPLICANT: KODAMA, YUKIO
 ; APPLICANT: FUJIMURA, TOMOKO
 ; APPLICANT: ASHIKARI, TOSHIHIKO
 ; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 ; FILE REFERENCE: 9-38-285
 ; CURRENT APPLICATION NUMBER: US/11/217,529
 ; CURRENT FILING DATE: 2005-09-02
 ; PRIOR APPLICATION NUMBER: US 10/932,182
 ; PRIOR FILING DATE: 2004-09-02
 ; NUMBER OF SEQ ID NOS: 197023
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 82353
 ; LENGTH: 606
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces pastorianus
 US-11-217-529-82353

```

Query Match      74.7%; Score 14.2; DB 7; Length 606;
Best Local Similarity 84.2%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTG 19
    |||||
Db 255 CTTTGGTCATCGCTCTTG 273

```

```

RESULT 26
US-11-217-529-1071/c
; Sequence 1071, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODANA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1071
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1071

```

```

Query Match          74.7%;   Score 14.2;   DB 7;   Length 771;
Best Local Similarity 84.2%;   Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 CTTTACTTCATAGCTTTG 19
    |||||

```



```
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7897
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-7897

Query Match          74.7%; Score 14.2; DB 6; Length 1333;
Best Local Similarity 84.2%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db 96 CTTTCTCTCAAAGGCTTTG 114

RESULT 32
US-11-285-701-17
; Sequence 17, Application US/11285701
; Publication No. US20060117415A1
; GENERAL INFORMATION:
; APPLICANT: Keenan, Ulrich
; APPLICANT: Linemann, Ute
; APPLICANT: Herbbers, Karin
; APPLICANT: Hillebrand, Helke
; TITLE OF INVENTION: EXPRESSION CASSETTES FOR ROOT-PREFERENTIAL EXPRESSION IN PLANTS
; FILE REFERENCE: 13173-00018-US
; CURRENT APPLICATION NUMBER: US/11/285,701
; CURRENT FILING DATE: 2005-11-22
; PRIOR APPLICATION NUMBER: EP 04028202.2
; PRIOR FILING DATE: 2004-11-27
; PRIOR APPLICATION NUMBER: EP 05002261.5
; PRIOR FILING DATE: 2005-02-03
; PRIOR APPLICATION NUMBER: EP 05002849.7
; PRIOR FILING DATE: 2005-02-11
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(1094)
; OTHER INFORMATION: coding for peroxidase ATP9a protein
US-11-285-701-17

Query Match          74.7%; Score 14.2; DB 7; Length 1333;
Best Local Similarity 84.2%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db 96 CTTTCTCTCAAAGGCTTTG 114

RESULT 33
US-11-217-529-780/c
; Sequence 780, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
```

```
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 780
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-780

Query Match          74.7%; Score 14.2; DB 7; Length 1566;
Best Local Similarity 84.2%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db 502 CTTCACTTGATAGTCTTTG 484

RESULT 34
US-11-217-529-77811/c
; Sequence 77811, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77811
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-77811

Query Match          74.7%; Score 14.2; DB 7; Length 1572;
Best Local Similarity 84.2%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTTTG 19
    ||||| ||||| ||||| |||||
Db 428 CTTTCTCTCATAGTCTTTG 410

RESULT 35
US-11-293-697-1270
; Sequence 1270, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1270
; LENGTH: 1860
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1270

Query Match      74.7%; Score 14.2; DB 7; Length 1860;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 1410 CTTTGCTTCATAGTATTGG 1428

RESULT 36
US-11-293-697-155
; Sequence 155, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 155
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-155

Query Match      74.7%; Score 14.2; DB 7; Length 1910;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 1652 CTTTACTTAATAGTATTGG 1670

RESULT 37
US-10-953-349-11687
; Sequence 11687, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCES-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11687
; LENGTH: 1986
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11687

Query Match      74.7%; Score 14.2; DB 6; Length 1986;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 133 CTTCAATTCATATCTTTG 151

RESULT 38
US-11-217-529-2941/c
; Sequence 2941, Application US/11217529
```

```
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2941
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2941

Query Match      74.7%; Score 14.2; DB 7; Length 2010;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 891 CTTTTCCTTCAAGTCCTTTG 873

RESULT 39
US-11-293-697-1439/c
; Sequence 1439, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1439
; LENGTH: 2594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1439

Query Match      74.7%; Score 14.2; DB 7; Length 2594;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
    ||||| ||||| ||||| |||||
Db 1719 CTTTCCTTTATAATCTTTG 1701

RESULT 40
US-11-217-529-78237/c
; Sequence 78237, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
```

; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78237
; LENGTH: 2628
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78237

Query Match 74.7%; Score 14.2; DB 7; Length 2628;
Best Local Similarity 84.2%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
||| ||||| |||||
Db 2148 CTTGGCTTCAATAATCTTTG 2130

RESULT 41
US-11-327-357-1/c
; Sequence 1, Application US/11327357
; Publication No. US20060105983A1
; GENERAL INFORMATION:
; APPLICANT: BTG International Ltd
; APPLICANT: BEARD DR, PETER
; APPLICANT: RAJ DR, KENNETH
; TITLE OF INVENTION: CYTOTOXIC AGENTS
; FILE REFERENCE: 142184W0
; CURRENT APPLICATION NUMBER: US/11/327,357
; CURRENT FILING DATE: 2006-01-09
; PRIOR APPLICATION NUMBER: 0009887.1
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4675
; TYPE: DNA
; ORGANISM: adeno-associated virus 2
US-11-327-357-1

Query Match 74.7%; Score 14.2; DB 7; Length 4675;
Best Local Similarity 84.2%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTTTG 19
||| ||||| |||||
Db 3793 CTTTCTTCATCGCTCTTG 3775

RESULT 42
US-11-217-529-138/c
; Sequence 138, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHKARI, TOSHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 138
; LENGTH: 936

; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-138

Query Match 73.7%; Score 14; DB 7; Length 936;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CTTCATAGTCCTTTG 19
||| ||||| |||||
Db 171 CTTCATAGTCCTTTG 158

RESULT 43
US-10-953-349-2188/c
; Sequence 2188, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2188
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2188

Query Match 73.7%; Score 14; DB 6; Length 1422;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CTTCATAGTCCTTTG 19
||| ||||| |||||
Db 803 CTTCATAGTCCTTTG 790

RESULT 44
US-10-953-349-8157/c
; Sequence 8157, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8157
; LENGTH: 2013
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8157

Query Match 73.7%; Score 14; DB 6; Length 2013;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CTTCATAGTCCTTTG 19
||| ||||| |||||
Db 809 CTTCATAGTCCTTTG 796

RESULT 45
US-11-217-529-190653
; Sequence 190653, Application US/11217529
; Publication No. US20060099612A1

```
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/217,529
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 190653
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-190653

Query Match      72.6%; Score 13.8; DB 7; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
   ||||| ||||| |||||
Db 4 TTCCTTTATAGTCTTTG 20

RESULT 46
US-10-471-571A-751/c
; Sequence 751, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471.571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 751
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-751

Query Match      72.6%; Score 13.8; DB 6; Length 183;
Best Local Similarity 88.2%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCTTT 18
   ||||| ||||| |||||
Db 83 TTTACTTCATAGTCTTT 67

RESULT 47
US-10-473-173-164/c
; Sequence 164, Application US/10473173
; Publication No. US2006008823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473.173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 164
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-164

Query Match      72.6%; Score 13.8; DB 6; Length 272;
Best Local Similarity 88.2%; Pred. No. 2.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
   ||||| ||||| ||||| |||||
Db 79 TTACATTATAGTCTTTG 63

RESULT 48
US-11-301-554-1379
; Sequence 1379, Application US/11301554
; Publication No. US20060088527A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C21
; CURRENT APPLICATION NUMBER: US/11/301.554
; CURRENT FILING DATE: 2005-12-13
; PRIOR APPLICATION NUMBER: US 10/283,017
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US 10/113,872
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/017,754
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902,941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849,626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736,457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702,705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677,419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671,325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658,824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1379
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-301-554-1379

Query Match      72.6%; Score 13.8; DB 7; Length 330;
Best Local Similarity 88.2%; Pred. No. 2.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
```


Db 96 TTACATTATAGTCTTTG 112
||||| |||||||

RESULT 49

US-11-217-529-174589
; Sequence 174589, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 174589
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-174589

Query Match 72.6%; Score 13.8; DB 7; Length 378;
Best Local Similarity 88.2%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
||| |||||||
Db 213 TTCTTTATAGTCTTTG 229

RESULT 50

US-10-488-619-1992/c
; Sequence 1992, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1992
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-1992

Query Match 72.6%; Score 13.8; DB 6; Length 564;
Best Local Similarity 88.2%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
||| |||||||
Db 479 TTCTTCATAGTCTTCG 463

RESULT 51

US-11-217-529-82746/c
; Sequence 82746, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; * APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO

; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82746
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82746

Query Match 72.6%; Score 13.8; DB 7; Length 696;
Best Local Similarity 88.2%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCTTT 18
||| |||||||
Db 602 TTAATTCATAGTCTTT 586

RESULT 52

US-11-217-529-77789/c
; Sequence 77789, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77789
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-77789

Query Match 72.6%; Score 13.8; DB 7; Length 729;
Best Local Similarity 88.2%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
||| |||||||
Db 454 TTATCATAGTCTTTG 438

RESULT 53

US-10-953-349-8745
; Sequence 8745, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252

```

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8745
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8745

```

Query Match	72.6%	Score 13.8;	DB 6;	Length 841;
Best Local Similarity	89.2%	Pred. No. 2.5e+02;		
Matches 15;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps	0;			

Qy 1 CTTTACTTCATAGTCTT 17
|||
Db 47 CTTTACTTCATAGTCTT 63

RESULT 54
US-11-217-529-166678/c
; Sequence 166678, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166678
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-166678

Query Match	72.6%	Score 13.8	DB 7	Length 867
Best Local Similarity	88.2%	Pred. No. 2.5e+02		
Matches 15	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

Qy 2 TTTACTTCATAGTCTTT 18
|||
db 845 TCTACTTCATAGCCTTT 829

```

RESULT 55
US-10-471-571A-3451/c
; Sequence 3451, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3451
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3451

```

Query Match	72.6%	Score 13.8;	DB 6;	Length 885;
Best Local Similarity	88.2%;	Pred. No. 2.5e+02;		
Matches 15;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

Qy 3 TTACTTCATAGTCTTTG 19
||| ||| ||| ||| |||
Db 793 TTAGTTCATAGTCTTGG 777

```

RESULT 56
US-10-471-571A-3035/c
; Sequence 3035, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3035
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3035

```

Query Match	72.6%	Score 13.8;	DB 6;	Length 900;
Best Local Similarity	89.2%;	Pred. No. 2.5e+02;		
Matches 15:	Conservative 2;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
|||||
D6 626 TTTACTTCATTTTCTTT 610

```

RESULT 57
US-10-953-349-7110
; Sequence 7110, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7110
; LENGTH: 946
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-7110

```

Query Match	72.6%;	Score 13.8;	DB 6;	Length 946;
Best Local Similarity	89.2%;	Pred. No. 2.5e+02;		
Matches 15:	Conservative	0;	Mismatches 2;	Indels 0;
	Gaps	0;		

Qy 1 CTTTACTTCATAGTCTT 17
|||
db 93 CTCTACTTCATATTCTT 109

```

RESULT 58
US-10-471-571A-3891/c
; Sequence 3891, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12

```


QY 2 TTTACTTCATAGTCTTT 18
|||||
Db 22 TTTACTTCAGAGTCTGT 38
|||||

RESULT 63

US-10-953-349-20194
; Sequence 20194, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20194
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20194

Query Match 72.6%; Score 13.8; DB 6; Length 1240;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 17
|||||
Db 592 CTTTACATCATATCTTT 608
|||||

RESULT 64

US-11-217-529-78180/c
; Sequence 78180, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78180
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78180

Query Match 72.6%; Score 13.8; DB 7; Length 1245;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 17
|||||
Db 132 CTTTACTTCATCTCTT 116
|||||

RESULT 65

US-10-953-349-1160
; Sequence 1160, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1160
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1160

Query Match 72.6%; Score 13.8; DB 6; Length 1272;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TTACTTCATAGTCTTTG 19
|||||
Db 1106 TTACTTCCTAGTCCTTG 1122
|||||

RESULT 66

US-10-953-349-20559/c
; Sequence 20559, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20559
; LENGTH: 1315
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20559

Query Match 72.6%; Score 13.8; DB 6; Length 1315;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTTACTTCATAGTCTTT 17
|||||
Db 1155 CTTTACTTCATGTATT 1139
|||||

RESULT 67

US-10-953-349-15758/c
; Sequence 15758, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15758
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-15758

Query Match 72.6%; Score 13.8; DB 6; Length 1332;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 2 TTTACTTCATAGTCTTT 18
      ||||| ||| |||||
Db 708 TTTACTTAATACTTT 692

RESULT 68
US-10-953-349-17823
; Sequence 17823, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17823
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-17823

Query Match 72.6%; Score 13.8; DB 6; Length 1333;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
      ||||| ||| |||||
Db 630 TATACTTCACAGTCTTT 646

RESULT 69
US-10-471-571A-4559
; Sequence 4559, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4559
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-4559

Query Match 72.6%; Score 13.8; DB 6; Length 1341;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
      ||||| ||| |||||
Db 371 TCTACTTCATAGTCGTT 387

RESULT 70
US-10-953-349-24191
; Sequence 24191, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
```

```
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24191
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-24191

Query Match 72.6%; Score 13.8; DB 6; Length 1404;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
      ||||| ||| |||||
Db 846 TATACTTCACAGTCTTT 862

RESULT 71
US-10-953-349-7778/c
; Sequence 7778, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7778
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-7778

Query Match 72.6%; Score 13.8; DB 6; Length 1420;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTT 17
      ||||| ||| |||||
Db 669 CTTTAATTCAAAGTCTT 653

RESULT 72
US-11-217-529-82109/c
; Sequence 82109, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82109
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82109

Query Match 72.6%; Score 13.8; DB 7; Length 1461;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      2 TTTACTTCATAGTCTTT 18
      ||||| ||||| |||||
Db      1415 TTTACATCATCGTCTTT 1399

RESULT 73
US-11-217-529-5036/c
; Sequence 5036, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5036
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5036

Query Match      72.6%; Score 13.8; DB 7; Length 1485;
Best Local Similarity 88.2%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
      ||||| ||||| |||||
Db      1424 TCTACTTCATAGTCTGT 1408

RESULT 74
US-11-217-529-82729/c
; Sequence 82729, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82729
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82729

Query Match      72.6%; Score 13.8; DB 7; Length 1605;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
      ||||| ||||| |||||
Db      692 TTTATTTCATGTCTTT 676
```

```
RESULT 75
US-11-217-529-421/c
; Sequence 421, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 421
; LENGTH: 1704
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-421

Query Match      72.6%; Score 13.8; DB 7; Length 1704;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
      ||||| ||||| |||||
Db      1107 TTTACTTCATAGTCTTT 1091

RESULT 76
US-11-217-529-4484/c
; Sequence 4484, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4484
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-4484

Query Match      72.6%; Score 13.8; DB 7; Length 1800;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 TTTACTTCATAGTCTTT 18
      ||||| ||||| |||||
Db      166 TTTCTTCATAGTATTT 150

RESULT 77
US-10-953-349-20342
; Sequence 20342, Application US/10953349
; Publication No. US20060107345A1
```

```
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20342
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20342

Query Match          72.6%; Score 13.8; DB 6; Length 1806;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCCTTT 18
   ||||| ||||| |||||
Db 1538 TATACCTCAGTCCTTT 1554

RESULT 78
US-10-953-349-1741
; Sequence 1741, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1741
; LENGTH: 1843
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1741

Query Match          72.6%; Score 13.8; DB 6; Length 1843;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19
   ||||| ||||| |||||
Db 1486 TTACTTGATATCTTTG 1502

RESULT 79
US-11-217-529-534/c
; Sequence 534, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 534
; LENGTH: 1854
```

```
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-534

Query Match          72.6%; Score 13.8; DB 7; Length 1854;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTACTTCATAGTCCTTT 18
   ||||| ||||| |||||
Db 1607 TTACTTCATATCTTT 1591

RESULT 80
US-10-953-349-18787
; Sequence 18787, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18787
; LENGTH: 1861
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-18787

Query Match          72.6%; Score 13.8; DB 6; Length 1861;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCCTT 17
   ||||| ||||| |||||
Db 109 CTTTCTTCATGTCTT 125

RESULT 81
US-10-471-571A-4139/c
; Sequence 4139, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4139
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-4139

Query Match          72.6%; Score 13.8; DB 6; Length 1899;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TTACTTCATAGTCCTTTG 19
   ||||| ||||| |||||
Db 1378 TTTCTTGATAGTCCTTG 1362

RESULT 82
US-11-217-529-75539/c
; Sequence 75539, Application US/11217529
```



```
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 75539
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-75539

Query Match      72.6%; Score 13.8; DB 7; Length 2004;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 TTACTTCATAGTCTTTG 19
Db      511 TTCTTCATATCTTTG 495

RESULT 83
US-11-217-529-565/c
; Sequence 565, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 565
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-565

Query Match      72.6%; Score 13.8; DB 7; Length 2031;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 TTACTTCATAGTCTTTG 19
Db      997 TTCTTCATATCTTTG 981

RESULT 84
US-10-471-571A-1409/c
; Sequence 1409, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
```

```
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1409
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-1409

Query Match      72.6%; Score 13.8; DB 6; Length 2190;
Best Local Similarity 88.2%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 TTTTACTTCATAGTCTTT 18
Db      1559 TTTTCTCATAGTCTTT 1543

RESULT 85
US-10-953-349-8153
; Sequence 8153, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8153
; LENGTH: 2320
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8153

Query Match      72.6%; Score 13.8; DB 6; Length 2320;
Best Local Similarity 88.2%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CTTTACTTCATAGTCTT 17
Db      102 CTTTACTTAATCGTCTT 118

RESULT 86
US-11-293-697-52/c
; Sequence 52, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 2577
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-52

Query Match      72.6%; Score 13.8; DB 7; Length 2577;
Best Local Similarity 88.2%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 2 TTTACTTCATAGTCTTT 18
||||| ||||| |||||
Db 1098 TTTACATCATAGTCTAT 1082

RESULT 87

US-11-293-697-749
; Sequence 749, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 749
; LENGTH: 2650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-749

Query Match 72.6%; Score 13.8; DB 7; Length 2650;
Best Local Similarity 88.2%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
||||| ||||| |||||
Db 742 TTTACTTCATCATCTTT 758

RESULT 88

US-11-293-697-2321/c
; Sequence 2321, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2321
; LENGTH: 2854
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-2321

Query Match 72.6%; Score 13.8; DB 7; Length 2854;
Best Local Similarity 88.2%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
||||| ||||| |||||
Db 2397 TTTACTTCAAAGTCATT 2381

RESULT 89

US-11-217-529-82302/c
; Sequence 82302, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO

; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82302
; LENGTH: 3102
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82302

Query Match 72.6%; Score 13.8; DB 7; Length 3102;
Best Local Similarity 88.2%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTTTACTTCATAGTCTT 17
||||| ||||| |||||
Db 2100 CTTTTCATAGTCTT 2084

RESULT 90

US-11-217-529-4879/c
; Sequence 4879, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4879
; LENGTH: 3273
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-4879

Query Match 72.6%; Score 13.8; DB 7; Length 3273;
Best Local Similarity 88.2%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
||||| ||||| |||||
Db 845 TCTACTTCATAGCCTTT 829

RESULT 91

US-11-217-529-77385/c
; Sequence 77385, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02

```
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77385
; LENGTH: 3447
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2669)..(2669)
; OTHER INFORMATION: a, c, g, t, unknown, or other
US-11-217-529-77385

Query Match      72.6%; Score 13.8; DB 7; Length 3447;
Best Local Similarity 88.2%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TTACTTCATAGTCCTT 18
Db      1123 TTAACCTCATAGTCCTT 1107
      ||| ||||| ||||| |||

RESULT 92
US-11-217-529-3930
; Sequence 3930, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3930
; LENGTH: 3528
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-3930

Query Match      72.6%; Score 13.8; DB 7; Length 3528;
Best Local Similarity 88.2%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 TTACTTCATAGTCCTTGG 19
Db      1776 TTCCTTTATAGTCCTTGG 1792
      ||| ||||| ||||| |||||

RESULT 93
US-10-953-349-2088/c
; Sequence 2088, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2088
; LENGTH: 4746
; TYPE: DNA

; ORGANISM: Arabidopsis thaliana
US-10-953-349-2088

Query Match      72.6%; Score 13.8; DB 6; Length 4746;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CTTTACTTCATAGTCCTT 17
Db      2284 CTTTACATCAGAGTCCTT 2268
      ||||| ||| ||||| |||||

RESULT 94
US-11-217-529-5476/c
; Sequence 5476, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5476
; LENGTH: 5283
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5476

Query Match      72.6%; Score 13.8; DB 7; Length 5283;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TTACTTCATAGTCCTT 18
Db      2918 TTTCCTTCATATCTTCTT 2902
      ||||| ||||| ||||| |||||

RESULT 95
US-10-506-549-3
; Sequence 3, Application US/10506549
; Publication No. US20060100417A1
; GENERAL INFORMATION:
; APPLICANT: APPLERA CORPORATION
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001361-US
; CURRENT APPLICATION NUMBER: US/10/506,549
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/361,343
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 394191
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(394191)
; OTHER INFORMATION: n = A,T,C or G
US-10-506-549-3

Query Match      72.6%; Score 13.8; DB 6; Length 394191;
```

Best Local Similarity 88.2%; Pred. NO. 2.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCTTT 18
||| ||||| |||||
Db 361068 TTGACTTCAAAGTCTTT 361084

RESULT 96

US-11-217-529-74966
; Sequence 74966, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 74966
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-74966

Query Match 70.5%; Score 13.4; DB 7; Length 25;
Best Local Similarity 93.3%; Pred. NO. 2.6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTTACTTCATAGTCT 16
||| ||||| |||||
Db 3 TTTACTTCACAGTCT 17

RESULT 97

US-10-471-571A-1517/c
; Sequence 1517, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1517
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-1517

Query Match 70.5%; Score 13.4; DB 6; Length 552;
Best Local Similarity 93.3%; Pred. NO. 3.6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TACTTCATAGTCTTT 18
||| ||||| |||||
Db 408 TACTTCATGCTCTTT 394

RESULT 98

US-10-953-349-15823/c

; Sequence 15823, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15823
; LENGTH: 878
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-15823

Query Match 70.5%; Score 13.4; DB 6; Length 878;
Best Local Similarity 93.3%; Pred. NO. 3.8e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TACTTCATAGTCTTT 18
||| ||||| |||||
Db 336 TACTTCATGCTTTT 322

RESULT 99

US-11-217-529-269/c
; Sequence 269, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 269
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-269

Query Match 70.5%; Score 13.4; DB 7; Length 966;
Best Local Similarity 93.3%; Pred. NO. 3.9e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ACTTCATAGTCTTTG 19
||| ||||| |||||
Db 863 ACTTCATGCTTTTG 849

RESULT 100

US-11-253-300-6
; Sequence 6, Application US/11253300
; Publication No. US20060107349A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, JIYI
; APPLICANT: WANG, ZENGJU
; TITLE OF INVENTION: INCREASED WAX PRODUCTION IN PLANTS
; FILE REFERENCE: NELE:045US
; CURRENT APPLICATION NUMBER: US/11/253,300
; CURRENT FILING DATE: 2005-10-18
; PRIOR APPLICATION NUMBER: 60/619,801
; PRIOR FILING DATE: 2004-10-18

•
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Medicago truncatula
US-11-253-300-6

Query Match 70.5%; Score 13.4; DB 7; Length 1023;
Best Local Similarity 93.3%; Pred. No. 3.9e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ACTTCATAGCTTTG 19
||| ||||| |||||
Db 200 ACTTCATAGGCTTTG 214

Search completed: June 10, 2006, 15:37:40
Job time : 17.1833 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:18:44 ; Search time 1145.2 Seconds
(without alignments)
2345.260 Million cell updates/sec

Title: US-10-600-816-33

Perfect score: 33.2

Sequence: 1 gcncaaycgtgcctcnc.....ayaargaytaygargttnaar 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	90.4	603	2	BD233463 Human pro
2	30	90.4	1212	2	BD209699 Compositi
3	30	90.4	1212	2	AK341505 Sequence
4	30	90.4	1619	2	AX078375 Sequence
5	30	90.4	1718	2	BD233473 Human pro
6	30	90.4	1826	5	AK172760 Homo sapi
7	30	90.4	2290	2	CQ723177 Sequence
8	30	90.4	2296	5	BC003665 Homo sapi
9	30	90.4	2297	5	AF506289 Homo sapi
10	30	90.4	2302	2	AX930411 Sequence
11	30	90.4	2302	5	AF095448 Homo sapi
12	30	90.4	2446	2	BD156680 Primer fo
13	30	90.4	2446	2	AX877483 Sequence
14	30	90.4	2446	5	AK001761 Homo sapi
15	30	90.4	2456	2	CQ981495 Sequence
16	30	90.4	2456	2	DD210040 Methods o
17	30	90.4	2456	2	AX549168 Sequence
18	30	90.4	3057	5	AK123672 Homo sapi

19	30	90.4	3371	2	AX188348 Sequence
20	30	90.4	6730	2	CQ894732 Sequence
c 21	30	90.4	161577	5	AC007688 Homo sapi
22	27.4	82.5	680	2	BD150346 Primer fo
23	27.4	82.5	680	2	AX870284 Sequence
24	22.2	66.9	224653	12	AC149246 Otolaur
c 25	22	66.3	179391	5	AC012558 Homo sapi
26	22	66.3	202050	12	AC060229 Homo sapi
c 27	22	66.3	222810	5	AC018977 Homo sapi
28	21.6	65.1	179591	5	AC093902 Homo sapi
c 29	21.6	65.1	182332	6	AL606745 Mouse DNA
c 30	21.6	65.1	203317	12	AC021062 Mus muscu
31	21.4	64.5	859	4	AK111295 Oryza sat
32	21.4	64.5	21062	13	CEV10G11R Caenorhab
33	21.4	64.5	110000	4	AP008217_262 Continuation (263
34	21.4	64.5	133581	4	AC134924 Oryza sat
35	21.4	64.5	164168	5	AL135927 Homo DNA
c 36	21.4	64.5	164179	5	AC007227 Homo sapi
37	21.4	64.5	186979	12	AC158014 Bos tauru
38	21.4	64.5	196244	12	AC152284 Bos tauru
39	21.2	63.9	1604	15	AF110497 Bartonell
40	21.2	63.9	110000	15	EX897700_01 Continuation (2 of
41	21.2	63.9	110000	15	EX897700_02 Continuation (3 of
c 42	21.2	63.9	203946	12	AC069465 Mus muscu
c 43	21.2	63.9	235241	6	AL928893 Mouse DNA
c 44	21	63.3	98142	12	AC159685 Bos tauru
c 45	21	63.3	142190	5	AC066598 Homo sapi
46	21	63.3	151830	12	AC024146 Homo sapi
c 47	21	63.3	152580	5	AC024164 Homo sapi
48	21	63.3	154010	12	AC024219 Homo sapi
c 49	21	63.3	157590	12	AC172585 Bos tauru
50	21	63.3	165646	12	CT027633 Danio rer
51	21	63.3	168908	11	CR361547 Zebrafish
c 52	21	63.3	178752	5	AC021889 Homo sapi
c 53	21	63.3	181736	5	AC087589 Homo sapi
54	21	63.3	185496	5	AC018832 Homo sapi
c 55	21	63.3	191265	12	AC068996 Homo sapi
56	21	63.3	197755	12	CT573362 Mus muscu
57	21	63.3	215523	12	CT573362 Bos tauru
c 58	21	63.3	228141	6	AC154912 Mus muscu
59	21	63.3	229726	12	AC095735 Rattus no
c 60	21	63.3	235642	6	CT010441 Mouse DNA
61	21	63.3	246535	12	AC109060 Rattus no
c 62	20.8	62.7	571	7	BV053582 S208P6041
63	20.8	62.7	3378	11	BC078095 Xenopus l
64	20.8	62.7	81744	11	EX927176 Zebrafish
c 65	20.8	62.7	110000	15	AB014295_15 o
66	20.8	62.7	139509	12	CR354544 Continuation (16 o
c 67	20.8	62.7	154959	6	AL662881 Mouse DNA
c 68	20.8	62.7	165987	11	AC147883 Xenopus t
c 69	20.8	62.7	171978	6	AC131779 Mus muscu
c 70	20.8	62.7	174929	12	AC162371 Mus muscu
71	20.8	62.7	194709	12	AC132640 Rattus no
72	20.8	62.7	195892	12	AC126211 Rattus no
c 73	20.8	62.7	210916	12	CT573319 Danio rer
c 74	20.8	62.7	212324	12	AC096459 Rattus no
c 75	20.8	62.7	219429	12	AC137342 Rattus no
76	20.8	62.7	232234	12	AC108634 Rattus no
c 77	20.8	62.7	240216	12	AC107096 Rattus no
78	20.8	62.7	241557	12	AC098000 Rattus no
79	20.8	62.7	243005	12	AC134740 Rattus no
80	20.8	62.7	287291	12	AC115384 Rattus no
81	20.8	62.7	319548	12	AC110863 Rattus no
c 82	20.8	62.7	349980	2	AX492783 Sequence
c 83	20.8	62.7	349980	2	AX553950 Sequence
84	20.6	62.0	770	7	BV639777 S217P6046
85	20.6	62.0	9060	15	AY615358 Mycoplasma
86	20.6	62.0	9720	15	MYCPMGAB L28424
87	20.6	62.0	44308	12	AC100253 Mus muscu
c 88	20.6	62.0	84234	5	AC107975 Homo sapi
89	20.6	62.0	110000	15	AB015450_04 Continuation (5 of
c 90	20.6	62.0	112585	12	AM050691 Triticum
c 91	20.6	62.0	142445	12	AC146668 Homo sapi

C 92	20.6	62.0	154471	5	AC022523	AC022523 Homo sapi	165	20.2	60.8	209476	12	AC021027	AC021027 Homo sapi
C 93	20.6	62.0	165165	12	CT025647	CT025647 Danio rer	C 166	20.2	60.8	213756	12	AC164200	AC164200 Bos tauru
C 94	20.6	62.0	171116	12	AP003651	AP003651 Mus muscu	C 167	20.2	60.8	214870	6	AL606496	AL606496 Mouse DNA
C 95	20.6	62.0	174250	6	AC164162	AC164162 Mus muscu	C 168	20.2	60.8	215467	12	AC013420	AC013420 Homo sapi
C 96	20.6	62.0	186376	6	CT027693	CT027693 Mouse DNA	C 169	20.2	60.8	219854	11	AC140791	AC140791 Gallus ga
C 97	20.6	62.0	191693	6	AC123073	AC123073 Mus muscu	C 170	20.2	60.8	228115	12	AC128376	AC128376 Rattus no
C 98	20.6	62.0	193668	6	AC147235	AC147235 Mus muscu	C 171	20.2	60.8	230020	12	AC106534	AC106534 Rattus no
C 99	20.6	62.0	207858	12	AC127132	AC127132 Rattus no	C 172	20.2	60.8	233440	12	AC094560	AC094560 Rattus no
C 100	20.6	62.0	219121	12	AC134718	AC134718 Rattus no	C 173	20.2	60.8	239385	12	AC109686	AC109686 Rattus no
C 101	20.6	62.0	221744	12	AC172135	AC172135 Bos tauru	C 174	20.2	60.8	248669	12	AC127097	AC127097 Rattus no
C 102	20.6	62.0	228501	6	AC094784	AC094784 Rattus no	C 175	20.2	60.8	250600	12	AC136646	AC136646 Rattus no
C 103	20.6	62.0	236068	12	AC121712	AC121712 Rattus no	C 176	20.2	60.8	254885	12	AC097132	AC097132 Rattus no
C 104	20.6	62.0	253387	12	AC133379	AC133379 Rattus no	C 177	20.2	60.8	256395	12	AC097025	AC097025 Rattus no
C 105	20.4	61.4	753	7	BV211953	BV211953 ONMS232 T	C 178	20.2	60.8	260290	12	AC094332	AC094332 Rattus no
C 106	20.4	61.4	1053	2	AR764489	AR764489 Sequence	C 179	20.2	60.8	270023	12	AC094829	AC094829 Rattus no
C 107	20.4	61.4	1074	2	BD162796	BD162796 Novel pol	C 180	20.2	60.8	271603	12	AC106318	AC106318 Rattus no
C 108	20.4	61.4	1074	2	AX120679	AX120679 Sequence	C 181	20.2	60.8	277252	12	AC175278	AC175278 Bos tauru
C 109	20.4	61.4	1201	2	AX773565	AX773565 Sequence	C 182	20.2	60.8	349980	2	CQ870195	CQ870195 Sequence
C 110	20.4	61.4	1857	13	DQ211932	DQ211932 Toxoplasma	C 183	20.2	60.2	282	7	AB133119	AB133119 Homo sapi
C 111	20.4	61.4	2045	4	SCU15217	UI5217 Saccharomyc	C 184	20.2	60.2	669	2	AX349817	AX349817 Sequence
C 112	20.4	61.4	2780	2	AX536762	AX536762 Sequence	C 185	20.2	60.2	718	7	BV611899	BV611899 S217P6209
C 113	20.4	61.4	3239	4	SCYOR267C	275175 S. cerevisia	C 186	20.2	60.2	2026	4	AK064556	AK064556 Oryza sat
C 114	20.4	61.4	72325	13	AC025722	AC025722 Caenorhab	C 187	20.2	60.2	3806	4	EN1269479	EN1269479 Emericell
C 115	20.4	61.4	106038	5	AC106747	AC106747 Homo sapi	C 188	20.2	60.2	12548	15	AE002165	AE002165 Chlamydog
C 116	20.4	61.4	110000	4	BA000034_1	Continuation (2 of	C 189	20.2	60.2	68890	5	AL662851	AL662851 Human DNA
C 117	20.4	61.4	110000	4	BA000034_3	Continuation (4 of	C 190	20.2	60.2	68890	5	AC004833	AC004833 Homo sapi
C 118	20.4	61.4	110000	4	AP008211_252	Continuation (253	C 191	20.2	60.2	80181	12	AC162649	AC162649 Bos tauru
C 119	20.4	61.4	110000	12	CR450853_0	CR450853 Danio rer	C 192	20.2	60.2	89319	12	AL929255	AL929255 Homo sapi
C 120	20.4	61.4	110000	15	CR522870_21	Continuation (22 o	C 193	20.2	60.2	107415	5	AL513365	AL513365 Human DNA
C 121	20.4	61.4	110000	15	BA000036_05	Continuation (6 of	C 194	20.2	60.2	110000	2	AR310754_08	Continuation (9 of
C 122	20.4	61.4	120891	12	AC011285	AC011285 Homo sapi	C 195	20.2	60.2	110000	2	AR607478_08	Continuation (9 of
C 123	20.4	61.4	146712	4	AC130604	AC130604 Oryza sat	C 196	20.2	60.2	110000	4	AP008208_064	Continuation (65 o
C 124	20.4	61.4	152945	12	AC130606	AC130606 Oryza sat	C 197	20.2	60.2	110000	12	CT005244_3	Continuation (9 of
C 125	20.4	61.4	157572	5	AL513210	AL513210 Human DNA	C 198	20.2	60.2	110000	15	AE001363_08	Continuation (9 of
C 126	20.4	61.4	161354	12	AC113374	AC113374 Homo sapi	C 199	20.2	60.2	110000	15	BA000008_08	Continuation (9 of
C 127	20.4	61.4	167905	5	AC116606	AC116606 Homo sapi	C 200	20.2	60.2	110752	11	BX649506	Continuation (9 of
C 128	20.4	61.4	174884	11	BX296542	BX296542 Zebrafish	C 201	20.2	60.2	114761	12	AC181823	Continuation (9 of
C 129	20.4	61.4	177857	12	AC144030	AC144030 Macaca mu	C 202	20.2	60.2	118572	5	AL133408	Continuation (9 of
C 130	20.4	61.4	187845	12	AC118983	AC118983 Rattus no	C 203	20.2	60.2	128729	5	AC137590	Continuation (9 of
C 131	20.4	61.4	203830	12	AC171610	AC171610 Bos tauru	C 204	20.2	60.2	131538	4	AP005846	Continuation (9 of
C 132	20.4	61.4	219390	12	AC114870	AC114870 Rattus no	C 205	20.2	60.2	137445	12	AP004780	Continuation (9 of
C 133	20.4	61.4	221270	12	AC123310	AC123310 Rattus no	C 206	20.2	60.2	150221	5	AC004593	Continuation (9 of
C 134	20.4	61.4	222741	12	AC128894	AC128894 Rattus no	C 207	20.2	60.2	156191	5	AC027248	Continuation (9 of
C 135	20.4	61.4	225487	6	AC162441	AC162441 Mus muscu	C 208	20.2	60.2	160250	12	AC174408	Continuation (9 of
C 136	20.4	61.4	226899	12	AC123119	AC123119 Rattus no	C 209	20.2	60.2	161451	5	AL355989	Continuation (9 of
C 137	20.4	61.4	235135	11	BX649266	BX649266 Zebrafish	C 210	20.2	60.2	170106	6	AC166995	Continuation (9 of
C 138	20.4	61.4	237432	12	AC123182	AC123182 Rattus no	C 211	20.2	60.2	170795	12	AC021346	Continuation (9 of
C 139	20.4	61.4	245486	12	AC125575	AC125575 Rattus no	C 212	20.2	60.2	170984	5	AL445312	Continuation (9 of
C 140	20.4	61.4	246499	12	AC140754	AC140754 Rattus no	C 213	20.2	60.2	171272	5	AC146225	Continuation (9 of
C 141	20.4	61.4	251382	12	AC162260	AC162260 Bos tauru	C 214	20.2	60.2	173239	5	AC016399	Continuation (9 of
C 142	20.4	61.4	263879	12	AC096060	AC096060 Rattus no	C 215	20.2	60.2	175086	12	AC139609	Continuation (9 of
C 143	20.4	61.4	266671	12	AC096126	AC096126 Rattus no	C 216	20.2	60.2	177206	11	CR354558	Continuation (9 of
C 144	20.4	61.4	293904	12	AC165519	AC165519 Bos tauru	C 217	20.2	60.2	179603	2	CQ870485	Continuation (9 of
C 145	20.4	61.4	349887	15	BX527749	BX527749 Corynebac	C 218	20.2	60.2	183540	12	AC026114	Continuation (9 of
C 146	20.4	61.4	349980	15	AX127144	AX127144 Sequence	C 219	20.2	60.2	185750	12	AC096875	Continuation (9 of
C 147	20.2	60.8	428	11	AY844405	AY844405 Hyla cavi	C 220	20.2	60.2	188768	11	BX908760	Continuation (9 of
C 148	20.2	60.8	52539	5	AL627447	AL627447 Human DNA	C 221	20.2	60.2	188972	5	AC020611	Continuation (9 of
C 149	20.2	60.8	110000	12	AL359456_3	Continuation (4 of	C 222	20.2	60.2	189791	12	AC152354	Continuation (9 of
C 150	20.2	60.8	110000	15	CR628336_22	Continuation (23 o	C 223	20.2	60.2	190762	12	AC163191	Continuation (9 of
C 151	20.2	60.8	110000	15	CR628337_22	Continuation (23 o	C 224	20.2	60.2	193375	12	AC172341	Continuation (9 of
C 152	20.2	60.8	110000	15	AE017354_22	Continuation (23 o	C 225	20.2	60.2	193551	11	BX649283	Continuation (9 of
C 153	20.2	60.8	129835	12	AC177867	AC177867 Zea mays	C 226	20.2	60.2	198254	12	AC147089	Continuation (9 of
C 154	20.2	60.8	130808	11	CR392029	CR392029 Zebrafish	C 227	20.2	60.2	199367	12	AC160727	Continuation (9 of
C 155	20.2	60.8	142116	12	AC157612	AC157612 Ornithorh	C 228	20.2	60.2	199688	12	AC163957	Continuation (9 of
C 156	20.2	60.8	143579	12	AC157845	AC157845 Ornithorh	C 229	20.2	60.2	204328	12	AL591119	Continuation (9 of
C 157	20.2	60.8	145729	5	AC117489	AC117489 Homo sapi	C 230	20.2	60.2	208057	5	AC146508	Continuation (9 of
C 158	20.2	60.8	168729	12	AC087418	AC087418 Homo sapi	C 231	20.2	60.2	208764	12	AC095320	Continuation (9 of
C 159	20.2	60.8	172412	12	AC134351	AC134351 Rattus no	C 232	20.2	60.2	209228	11	BX890613	Continuation (9 of
C 160	20.2	60.8	172770	12	AC148425	AC148425 Ornithorh	C 233	20.2	60.2	209921	12	AC163135	Continuation (9 of
C 161	20.2	60.8	182439	12	AL607037	AL607037 Homo sapi	C 234	20.2	60.2	212371	12	CR936381	Continuation (9 of
C 162	20.2	60.8	182639	5	AC104307	AC104307 Homo sapi	C 235	20.2	60.2	212429	11	BX571828	Continuation (9 of
C 163	20.2	60.8	187159	5	AC090941	AC090941 Homo sapi	C 236	20.2	60.2	236484	11	BX539341	Continuation (9 of
C 164	20.2	60.8	205754	12	AC097001	AC097001 Rattus no	C 237	20.2	60.2	240052	12	AC169766	Continuation (9 of

238	20	60.2	30066	15	AE017159	AE017159 Chlamydom	311	19.8	59.6	245698	12	AC124841	AC124841 Rattus no
239	20	60.2	317233	12	AC130067	AC130067 Rattus no	312	19.8	59.6	246029	12	AC168577	AC168577 Strongylo
240	20	60.2	349980	2	C0870484	C0870484 Sequence	c 313	19.8	59.6	246613	12	AC126851	AC126851 Rattus no
241	19.8	59.6	289	7	G89085	G89085 S208f6370FC	c 314	19.8	59.6	253631	12	AC126155	AC126155 Rattus no
242	19.8	59.6	428	11	AY844377	AY844377 Duellmano	c 315	19.8	59.6	256998	12	AC115751	AC115751 Mus muscu
243	19.8	59.6	428	11	AY844388	AY844388 Hyla anne	c 316	19.8	59.6	259730	12	AC094880	AC094880 Rattus no
244	19.8	59.6	428	11	AY844389	AY844389 Hyla arbo	c 317	19.8	59.6	262632	12	AC094326	AC094326 Rattus no
245	19.8	59.6	428	11	AY844391	AY844391 Hyla aren	c 318	19.8	59.6	275706	12	AC120591	AC120591 Rattus no
246	19.8	59.6	428	11	AY844420	AY844420 Hyla japo	c 319	19.8	59.6	293854	12	AC126646	AC126646 Rattus no
247	19.8	59.6	428	11	AY844462	AY844462 Hyla equi	c 320	19.8	59.6	299420	12	AC094449	AC094449 Rattus no
248	19.8	59.6	428	11	AY844465	AY844465 Hyla vers	c 321	19.8	59.6	299876	12	AC106196	AC106196 Rattus no
249	19.8	59.6	428	11	AY844466	AY844466 Hyla walk	c 322	19.8	59.6	340000	5	AP001684	AP001684 Homo sapi
250	19.8	59.6	428	11	AY844518	AY844518 Scinax el	c 323	19.6	59.0	356	2	CS066201	CS066201 Sequence
251	19.8	59.6	428	11	AY844519	AY844519 Scinax fu	c 324	19.6	59.0	428	11	AY844358	AY844358 Acris cre
252	19.8	59.6	428	11	AY844537	AY844537 Chigloss	c 325	19.6	59.0	428	11	AY844359	AY844359 Acris gry
253	19.8	59.6	1512	11	AY571662	AY571662 Hyla meri	c 326	19.6	59.0	428	11	AY844434	AY844434 Hyla mixe
254	19.8	59.6	1512	11	AY583339	AY583339 Hyla meri	c 327	19.6	59.0	1005	2	AR547935	AR547935 Sequence
255	19.8	59.6	1855	4	AK072046	AK072046 Oryza sat	c 328	19.6	59.0	1423	4	BT013914	BT013914 Lycopersi
256	19.8	59.6	2002	4	AK105688	AK105688 Oryza sat	c 329	19.6	59.0	2300	4	AF253056	AF253056 Candida a
257	19.8	59.6	2023	4	AK056632	AK056632 Oryza sat	c 330	19.6	59.0	2454	15	AB078638	AB078638 Helicobac
258	19.8	59.6	2500	4	CCY11081	Y11081 C. cinereus	c 331	19.6	59.0	26359	5	AL512268	AL512268 Human DNA
259	19.8	59.6	16745	6	AY257203	AY257203 Mus muscu	c 332	19.6	59.0	62288	5	AC005571	AC005571 Homo sapi
260	19.8	59.6	44455	4	YSCP9513	Y10333 Saccharomyc	c 333	19.6	59.0	68229	5	AC007090	AC007090 Homo sapi
261	19.8	59.6	51621	5	AP000657	AP000657 Homo sapi	c 334	19.6	59.0	94001	13	AC171141	AC171141 Helobdell
262	19.8	59.6	64829	12	AC173369	AC173369 Strongylo	c 335	19.6	59.0	98399	5	AL162291	AL162291 Human DNA
263	19.8	59.6	71927	12	AC113149	AC113149 Homo sapi	c 336	19.6	59.0	100000	15	BA000011_02	BA000011_02
264	19.8	59.6	73442	12	AC1130692	AC1130692 Homo sapi	c 337	19.6	59.0	100000	5	AB020865	AB020865 Homo sapi
265	19.8	59.6	93409	12	CT573813	CT573813 Danio rer	c 338	19.6	59.0	101377	12	AC152053	AC152053 Medicago
266	19.8	59.6	104374	11	EX649313	EX649313 Zebrafish	c 339	19.6	59.0	109739	12	AP008212_196	AP008212_196
267	19.8	59.6	110000	4	AP008212_110	Continuation (111	c 340	19.6	59.0	110000	15	CR522870_33	CR522870_33
268	19.8	59.6	110000	13	AC164548_2	Continuation (3 of	c 341	19.6	59.0	11446	4	AC143341	AC143341 Medicago
269	19.8	59.6	113812	4	AC149135	AC149135 Medicago	c 342	19.6	59.0	117329	4	AC148177	AC148177 Triticum
270	19.8	59.6	123291	5	AC104070	AC104070 Homo sapi	c 343	19.6	59.0	121991	12	AM050677	AM050677 Triticum
271	19.8	59.6	129206	12	HS03230H6	AL121940 Homo sapi	c 344	19.6	59.0	129756	5	HS875H10	HS875H10 Human DNA
272	19.8	59.6	142814	4	OSJN00253	AL731611 Oryza sat	c 345	19.6	59.0	136586	14	AC157604	AC157604 Ornithorh
273	19.8	59.6	143820	12	AC135843	AC135843 Felis cat	c 346	19.6	59.0	143722	4	AP005723	AP005723 Oryza sat
274	19.8	59.6	146559	5	AC013262	AC013262 Homo sapi	c 347	19.6	59.0	144712	12	AC123559	AC123559 Rattus no
275	19.8	59.6	155476	6	AC154816	AC154816 Mus muscu	c 348	19.6	59.0	150549	12	AP004794	AP004794 Oryza sat
276	19.8	59.6	157426	5	AC018449	AC018449 Homo sapi	c 349	19.6	59.0	153381	12	AP001990	AP001990 Homo sapi
277	19.8	59.6	159393	5	AC160568	AC160568 Pan trogl	c 350	19.6	59.0	156929	12	AC026854	AC026854 Homo sapi
278	19.8	59.6	161475	5	HS164C20	AL090929 Human DNA	c 351	19.6	59.0	157336	12	AC079787	AC079787 Homo sapi
279	19.8	59.6	161566	12	AC136462	AC136462 Rattus no	c 352	19.6	59.0	158794	12	AC023961	AC023961 Homo sapi
280	19.8	59.6	162977	12	AC177178	AC177178 Strongylo	c 353	19.6	59.0	159403	5	AC146155	AC146155 Pan trogl
281	19.8	59.6	162977	12	AC177178	AC177178 Strongylo	c 354	19.6	59.0	160835	5	AC087273	AC087273 Homo sapi
282	19.8	59.6	16285	12	AC174839	AC174839 Echinosps	c 355	19.6	59.0	161194	12	AC018945	AC018945 Homo sapi
283	19.8	59.6	170364	12	AC178856	AC178856 Strongylo	c 356	19.6	59.0	161675	12	AP001975	AP001975 Homo sapi
284	19.8	59.6	171656	12	AC178941	AC178941 Strongylo	c 357	19.6	59.0	162515	5	AC093206	AC093206 Homo sapi
285	19.8	59.6	174540	6	AC119900	AC119900 Mus muscu	c 358	19.6	59.0	162515	5	AC107032	AC107032 Homo sapi
286	19.8	59.6	174772	12	AC168491	AC168491 Strongylo	c 359	19.6	59.0	162515	5	AC005510	AC005510 Homo sapi
287	19.8	59.6	175339	5	AC005772	AC005772 Homo sapi	c 360	19.6	59.0	168468	12	AC109295	AC109295 Mus muscu
288	19.8	59.6	175840	12	AC170828	AC170828 Bos tauru	c 361	19.6	59.0	173118	5	CT025565	CT025565 Mouse DNA
289	19.8	59.6	175859	12	AC161486	AC161486 Mus muscu	c 362	19.6	59.0	175523	6	AM050683	AM050683 Triticum
290	19.8	59.6	180851	12	AC178594	AC178594 Strongylo	c 363	19.6	59.0	176804	12	AC053484	AC053484 Homo sapi
291	19.8	59.6	180851	12	AC090210	AC090210 Homo sapi	c 364	19.6	59.0	177244	12	AC123288	AC123288 Rattus no
292	19.8	59.6	181211	12	AP001904	AP001904 Homo sapi	c 365	19.6	59.0	179680	12	AL669953	AL669953 Mouse DNA
293	19.8	59.6	183385	6	AC117206	AC117206 Mus muscu	c 366	19.6	59.0	179956	12	AC098465	AC098465 Rattus no
294	19.8	59.6	188083	6	AC127309	AC127309 Mus muscu	c 367	19.6	59.0	180924	6	AC164175	AC164175 Mus muscu
295	19.8	59.6	188616	12	AC156152	AC156152 Bos tauru	c 368	19.6	59.0	181220	6	AC136008	AC136008 Mus muscu
296	19.8	59.6	200728	12	AC131357	AC131357 Rattus no	c 369	19.6	59.0	185724	12	AC146285	AC146285 Callibebu
297	19.8	59.6	201237	12	AC111505	AC111505 Rattus no	c 370	19.6	59.0	186744	12	AC176383	AC176383 Strongylo
298	19.8	59.6	203441	12	AC167131	AC167131 Mus muscu	c 371	19.6	59.0	186901	12	AF267167	AF267167 Homo sapi
299	19.8	59.6	206246	12	AC132970	AC132970 Rattus no	c 372	19.6	59.0	190735	12	AC026084	AC026084 Homo sapi
300	19.8	59.6	207088	6	AC160635	AC160635 Mus muscu	c 373	19.6	59.0	192992	12	AC145957	AC145957 Gallus ga
301	19.8	59.6	211343	5	AC010928	AC010928 Homo sapi	c 374	19.6	59.0	193614	6	AC110219	AC110219 Mus muscu
302	19.8	59.6	213552	5	AC010928	AC010928 Homo sapi	c 375	19.6	59.0	193614	6	AC124157	AC124157 Equus cab
303	19.8	59.6	215020	12	AC170695	AC170695 Bos tauru	c 376	19.6	59.0	193986	14	AP006249	AP006249 Homo sapi
304	19.8	59.6	22032	12	AC096942	AC096942 Rattus no	c 377	19.6	59.0	195290	5	AL772387	AL772387 Mouse DNA
305	19.8	59.6	222745	12	AC127100	AC127100 Rattus no	c 378	19.6	59.0	195412	6	AC165082	AC165082 Mus muscu
306	19.8	59.6	224334	12	AC096041	AC096041 Rattus no	c 379	19.6	59.0	196132	6	AC138192	AC138192 Mus muscu
307	19.8	59.6	231639	12	AC179487	AC179487 Strongylo	c 380	19.6	59.0	209937	6	AC138192	AC138192 Mus muscu
308	19.8	59.6	234462	12	AC109738	AC109738 Rattus no	c 381	19.6	59.0	209937	6	AC138192	AC138192 Mus muscu
309	19.8	59.6	235452	12	AC128416	AC128416 Rattus no	c 382	19.6	59.0	209937	6	AC138192	AC138192 Mus muscu
310	19.8	59.6	235468	6	AC105586	AC105586 Rattus no	c 383	19.6	59.0	209937	6	AC138192	AC138192 Mus muscu
311	19.8	59.6	235468	6	CR392369	CR392369 Danio rer	c 384	19.6	59.0	209937	6	AC138192	AC138192 Mus muscu

C 384	19.6	59.0	209945	12	AC139756	AC139756 Mus muscu	C 457	19.4	58.4	164040	12	AC147530	AC147530 Ocolemur
C 385	19.6	59.0	211300	12	CR790363	CR790363 Danio rer	C 458	19.4	58.4	164175	12	AC119056	AC119056 Papio anu
C 386	19.6	59.0	211465	6	AL646046	AL646046 Mouse DNA	C 459	19.4	58.4	165496	5	AC104461	AC104461 Homo sapi
C 387	19.6	59.0	221115	12	AC0931242	AC0931242 Homo sapi	C 460	19.4	58.4	166118	14	CT009560	CT009560 Pig DNA e
C 388	19.6	59.0	221708	12	AC103447	AC103447 Rattus no	C 461	19.4	58.4	170979	12	AC118896	AC118896 Rattus no
C 389	19.6	59.0	221953	12	AC130078	AC130078 Rattus no	C 462	19.4	58.4	171585	12	AC148301	AC148301 Pan trogl
C 390	19.6	59.0	222373	12	AC129363	AC129363 Rattus no	C 463	19.4	58.4	171806	6	AC113510	AC113510 Mus muscu
C 391	19.6	59.0	226029	12	AC171007	AC171007 Gallus ga	C 464	19.4	58.4	171822	12	CT025616	CT025616 Mus muscu
C 392	19.6	59.0	230703	12	AC120663	AC120663 Rattus no	C 465	19.4	58.4	172835	12	AC121174	AC121174 Rattus no
C 393	19.6	59.0	233112	12	AC172238	AC172238 Bos tauru	C 466	19.4	58.4	174545	12	AC176158	AC176158 Strongylo
C 394	19.6	59.0	242151	12	AC159523	AC159523 Bos tauru	C 467	19.4	58.4	174582	12	AC118488	AC118488 Rattus no
C 395	19.6	59.0	247511	12	AC096432	AC096432 Rattus no	C 468	19.4	58.4	175485	5	AC147539	AC147539 Pan trogl
C 396	19.6	59.0	257513	12	AC127093	AC127093 Rattus no	C 469	19.4	58.4	175504	12	AC153161	AC153161 Glycine m
C 397	19.6	59.0	270120	12	AC128627	AC128627 Rattus no	C 470	19.4	58.4	176241	12	AC167936	AC167936 Sus scrof
C 398	19.6	59.0	271200	12	AC131538	AC131538 Rattus no	C 471	19.4	58.4	176861	5	AC097065	AC097065 Homo sapi
C 399	19.6	59.0	271233	12	AC096597	AC096597 Rattus no	C 472	19.4	58.4	178044	12	AC121663	AC121663 Rattus no
C 400	19.6	59.0	275292	12	AC129270	AC129270 Rattus no	C 473	19.4	58.4	178246	12	AC124162	AC124162 Rattus no
C 401	19.6	59.0	275443	12	AC096407	AC096407 Rattus no	C 474	19.4	58.4	178872	12	CT573008	CT573008 Danio rer
C 402	19.6	59.0	276163	12	AC130878	AC130878 Rattus no	C 475	19.4	58.4	180072	5	AC017092	AC017092 Homo sapi
C 403	19.6	59.0	283341	12	AC128766	AC128766 Rattus no	C 476	19.4	58.4	180150	5	AC118138	AC118138 Homo sapi
C 404	19.6	59.0	294915	12	AC155005	AC155005 Bos tauru	C 477	19.4	58.4	180161	12	AC180837	AC180837 Strongylo
C 405	19.6	59.0	295001	12	AC118175	AC118175 Rattus no	C 478	19.4	58.4	181234	12	AC167935	AC167935 Sus scrof
C 406	19.6	59.0	309322	12	AC119578	AC119578 Rattus no	C 479	19.4	58.4	182902	5	AC092405	AC092405 Papio anu
C 407	19.6	59.0	318862	12	AC109547	AC109547 Rattus no	C 480	19.4	58.4	183118	12	AC160722	AC160722 Aotus nan
C 408	19.4	58.4	312	2	C0657729	C0657729 Sequence	C 481	19.4	58.4	183412	12	AP001899	AP001899 Homo sapi
C 409	19.4	58.4	428	11	AY844487	AY844487 Osteopillu	C 482	19.4	58.4	184164	5	AC091020	AC091020 Homo sapi
C 410	19.4	58.4	559	7	G09062	G09062 human STS C	C 483	19.4	58.4	184951	6	AC155292	AC155292 Mus muscu
C 411	19.4	58.4	660	1	AY214920	AY214920 Unculture	C 484	19.4	58.4	185763	5	AF152363	AF152363 Homo sapi
C 412	19.4	58.4	987	11	AY543804	AY543804 Salmo sal	C 485	19.4	58.4	189108	12	AC178034	AC178034 Strongylo
C 413	19.4	58.4	992	2	C0872398	C0872398 Sequence	C 486	19.4	58.4	189208	11	CR847789	CR847789 Zebrafish
C 414	19.4	58.4	1056	7	BV176987	BV176987 sgsm91960	C 487	19.4	58.4	189281	12	AC012458	AC012458 Homo sapi
C 415	19.4	58.4	1957	11	AY789437S3	AY789439 Oncorhync	C 488	19.4	58.4	190363	5	AC000052	AC000052 Homo sapi
C 416	19.4	58.4	2259	5	AF410455	AF410455 Homo sapi	C 489	19.4	58.4	191041	12	AC024114	AC024114 Mus muscu
C 417	19.4	58.4	2321	2	AX086673	AX086673 Sequence	C 490	19.4	58.4	191434	5	AC087477	AC087477 Homo sapi
C 418	19.4	58.4	2322	5	BC019008	BC019008 Homo sapi	C 491	19.4	58.4	192653	5	AC124781	AC124781 Homo sapi
C 419	19.4	58.4	2324	5	BC007823	BC007823 Homo sapi	C 492	19.4	58.4	193015	5	AC008460	AC008460 Homo sapi
C 420	19.4	58.4	2416	2	AR578216	AR578216 Sequence	C 493	19.4	58.4	193456	12	AC149228	AC149228 Pan trogl
C 421	19.4	58.4	2737	4	AK111149	AK111149 Oryza sat	C 494	19.4	58.4	197483	5	AC104164	AC104164 Homo sapi
C 422	19.4	58.4	3738	2	AF376414	AF376414 Sequence	C 495	19.4	58.4	198488	5	AC023794	AC023794 Homo sapi
C 423	19.4	58.4	8271	5	AF489699	AF489699 Homo sapi	C 496	19.4	58.4	200061	6	AC153535	AC153535 Mus muscu
C 424	19.4	58.4	11936	15	AE005946	AE005946 Caulobact	C 497	19.4	58.4	200355	12	AC184743	AC184743 Bos tauru
C 425	19.4	58.4	15229	14	AM182498	AM182498 Equus cab	C 498	19.4	58.4	201376	12	BX470152	BX470152 Mus tauru
C 426	19.4	58.4	46077	12	AC137289	AC137289 Rattus no	C 499	19.4	58.4	202197	5	BS000130	BS000130 Pan trogl
C 427	19.4	58.4	58143	12	AC137177	AC137177 Rattus no	C 500	19.4	58.4	202412	6	AL670673	AL670673 Mouse DNA
C 428	19.4	58.4	79843	5	AL353664	AL353664 Human DNA	C 501	19.4	58.4	205118	12	AC118478	AC118478 Mus muscu
C 429	19.4	58.4	80589	11	BX470257	BX470257 Zebrafish	C 502	19.4	58.4	205174	5	AC011804	AC011804 Homo sapi
C 430	19.4	58.4	81896	5	AP001614	AP001614 Homo sapi	C 503	19.4	58.4	207961	12	AC111825	AC111825 Rattus no
C 431	19.4	58.4	90526	5	AC092393	AC092393 Homo sapi	C 504	19.4	58.4	211056	6	AL591207	AL591207 Mouse DNA
C 432	19.4	58.4	104980	12	CT030699	CT030699 Danio rer	C 505	19.4	58.4	212548	6	AC102372	AC102372 Mus muscu
C 433	19.4	58.4	110000	4	CR382139_08	Continuation (9 of	C 506	19.4	58.4	213033	12	AC131482	AC131482 Rattus no
C 434	19.4	58.4	110000	4	AP082029_268	Continuation (9 of	C 507	19.4	58.4	219053	12	AC167784	AC167784 Glycine m
C 435	19.4	58.4	110000	12	AC109413_0	Continuation (269	C 508	19.4	58.4	212556	5	AP001743	AP001743 Homo sapi
C 436	19.4	58.4	110000	12	AC112029_0	AC112029 Rattus no	C 509	19.4	58.4	222161	12	AC147935	AC147935 Ocolemur
C 437	19.4	58.4	110000	12	AC112029_1	Continuation (2 of	C 510	19.4	58.4	224058	5	AC008513	AC008513 Homo sapi
C 438	19.4	58.4	110000	15	CP000297_08	Continuation (9 of	C 511	19.4	58.4	225829	6	AC129609	AC129609 Mus muscu
C 439	19.4	58.4	118327	4	AC004521	AC004521 Arabidops	C 512	19.4	58.4	228117	12	AC170632	AC170632 Bos tauru
C 440	19.4	58.4	118362	12	AC168753	AC168753 Strongylo	C 513	19.4	58.4	234607	12	AC152095	AC152095 Bos tauru
C 441	19.4	58.4	121661	13	AC150385	AC150385 Branchios	C 514	19.4	58.4	236720	12	AC123083	AC123083 Rattus no
C 442	19.4	58.4	124706	11	BX927312	BX927312 Zebrafish	C 515	19.4	58.4	237400	12	AC137383	AC137383 Rattus no
C 443	19.4	58.4	135931	4	AC097936	AC097936 Oryza sat	C 516	19.4	58.4	239242	12	AC099428	AC099428 Strongylo
C 444	19.4	58.4	141710	12	AC170079	AC170079 Loxodonta	C 517	19.4	58.4	240673	12	AC181744	AC181744 Rattus no
C 445	19.4	58.4	143527	6	AC115044	AC115044 Mus muscu	C 518	19.4	58.4	249433	12	AC051620	AC051620 Mus muscu
C 446	19.4	58.4	151207	6	AL604025	AL604025 Mouse DNA	C 519	19.4	58.4	260409	5	AC004019	AC004019 Homo sapi
C 447	19.4	58.4	152033	5	AC027319	AC027319 Homo sapi	C 520	19.4	58.4	261376	12	AC098996	AC098996 Rattus no
C 448	19.4	58.4	152413	12	AC181948	AC181948 Strongylo	C 521	19.4	58.4	265169	12	AC110476	AC110476 Rattus no
C 449	19.4	58.4	156207	12	AC025507	AC025507 Homo sapi	C 522	19.4	58.4	272400	12	AC110146	AC110146 Rattus no
C 450	19.4	58.4	157430	6	AC127678	AC127678 Mus muscu	C 523	19.4	58.4	276015	12	CR759962	CR759962 Homo sapi
C 451	19.4	58.4	157594	6	AC157928	AC157928 Mus muscu	C 524	19.4	58.4	276913	12	AC176363	AC176363 Strongylo
C 452	19.4	58.4	161005	11	BX537351	BX537351 Zebrafish	C 525	19.4	58.4	283299	12	AC112576	AC112576 Rattus no
C 453	19.4	58.4	161406	12	AC121326	AC121326 Homo sapi	C 526	19.4	58.4	287233	12	AC110639	AC110639 Rattus no
C 454	19.4	58.4	163009	5	AC011884	AC011884 Homo sapi	C 527	19.4	58.4	306522	12	AC137295	AC137295 Rattus no
C 455	19.4	58.4	163503	11	BX927205	BX927205 Zebrafish	C 528	19.4	58.4	349635	19.4	AC093978	AC093978 Rattus no
C 456	19.4	58.4	163612	12	AC011794	AC011794 Homo sapi	C 529	19.2	57.8	428	11	AY844413	AY844413 Hyla cyci

530	19.2	57.8	428	11	AY844438	AY844438 Hyla neph	c 603	19.2	57.8	31268	13	CEH08M01	293384 Caenorhabdi
c 531	19.2	57.8	430	7	BV292436	BV292436 S232P6431	604	19.2	57.8	31434	13	CEY68A4A	AL021503 Caenorhab
c 532	19.2	57.8	629	7	BV048212	BV048212 S212P6044	605	19.2	57.8	32009	13	CET06D8	249130 Caenorhabdi
c 533	19.2	57.8	664	11	AY609297	AY609297 Chlorospi	c 606	19.2	57.8	32488	13	AF100307	AF100307 Caenorhab
c 534	19.2	57.8	668	7	BV388502	BV388502 S244P617R	607	19.2	57.8	32591	13	CEFI16D3	278062 Caenorhabdi
c 535	19.2	57.8	673	11	AY609296	AY609296 Chlorospi	608	19.2	57.8	32838	13	AF003385	AF003385 Caenorhab
c 536	19.2	57.8	689	7	BV655500	BV655500 S216P6015	c 609	19.2	57.8	32960	13	CET27E5	282284 Caenorhabdi
c 537	19.2	57.8	829	11	AF109418	AF109418 Icterus m	c 610	19.2	57.8	33802	13	CER07H7	281107 Caenorhabdi
c 538	19.2	57.8	842	11	AF382968	AF382968 Seiurus a	c 611	19.2	57.8	34073	13	CEFI11D11	281500 Caenorhabdi
c 539	19.2	57.8	998	11	AF447354	AF447354 Seiurus a	c 612	19.2	57.8	34494	13	AC006633	AC006633 Caenorhab
c 540	19.2	57.8	1001	2	AK582469	AK582469 Sequence	c 613	19.2	57.8	34494	13	AC006633	AC006633 Caenorhab
c 541	19.2	57.8	1293	4	AK106009	AK106009 Oryza sat	c 614	19.2	57.8	35061	13	U47144	U47144 Caenorhabdi
c 542	19.2	57.8	1599	6	BC034063	BC034063 Mus muscu	c 615	19.2	57.8	35077	13	CEY38H8A	AL021483 Caenorhab
c 543	19.2	57.8	1670	11	AF468612	AF468612 Alimophila	c 616	19.2	57.8	35950	13	U49947	U49947 Caenorhabdi
c 544	19.2	57.8	1671	11	AF468611	AF468611 Alimophila	c 617	19.2	57.8	35950	13	U49947	U49947 Caenorhabdi
c 545	19.2	57.8	1704	13	AK115593	AK115593 Clona int	c 618	19.2	57.8	36170	13	CER10E4	250874 Caenorhabdi
c 546	19.2	57.8	1716	10	DQ228931	DQ228931 Newcastle	c 619	19.2	57.8	36170	13	U58728	U58728 Caenorhabdi
c 547	19.2	57.8	1716	10	DQ234586	DQ234586 Newcastle	c 620	19.2	57.8	36429	13	U29380	U29380 Caenorhabdi
c 548	19.2	57.8	1716	10	DQ234587	DQ234587 Newcastle	c 621	19.2	57.8	37427	13	CEH12119	AF039049 Caenorhab
c 549	19.2	57.8	1801	10	DQ023556	DQ023556 Newcastle	c 622	19.2	57.8	38375	13	CEY56D5	298851 Caenorhabdi
c 550	19.2	57.8	1828	2	AK698569	AK698569 Sequence	c 623	19.2	57.8	38861	13	CEC06B8	269662 Caenorhabdi
c 551	19.2	57.8	1828	2	AK338236	AK338236 Sequence	c 624	19.2	57.8	39333	13	CEB0564	281463 Caenorhabdi
c 552	19.2	57.8	1899	13	AF478698	AF478698 Drosophil	c 625	19.2	57.8	39902	13	CEB03H4	281492 Caenorhabdi
c 553	19.2	57.8	1899	13	AF478699	AF478699 Drosophil	c 626	19.2	57.8	40094	13	CEY44F4	237092 Caenorhabdi
c 554	19.2	57.8	1899	13	AF478700	AF478700 Drosophil	c 627	19.2	57.8	40145	13	CEC36A4	266495 Caenorhabdi
c 555	19.2	57.8	1936	6	BC036174	BC036174 Mus muscu	c 628	19.2	57.8	40641	13	AC006816	AC006816 Caenorhab
c 556	19.2	57.8	1971	13	AY084098	AY084098 Drosophil	c 629	19.2	57.8	41299	13	CEC06H5	292775 Caenorhabdi
c 557	19.2	57.8	2090	2	CQ588981	CQ588981 Sequence	c 630	19.2	57.8	41811	13	U58746	U58746 Caenorhabdi
c 558	19.2	57.8	2102	6	BC039217	BC039217 Mus muscu	c 631	19.2	57.8	42297	13	CEC33A12	268493 Caenorhabdi
c 559	19.2	57.8	2142	13	AF478697	AF478697 Drosophil	c 632	19.2	57.8	42300	13	U80452	U80452 Caenorhabdi
c 560	19.2	57.8	2166	13	AY071591	AY071591 Drosophil	c 633	19.2	57.8	42574	13	AF039053	AF039053 Caenorhab
c 561	19.2	57.8	2174	6	BC036173	BC036173 Mus muscu	c 634	19.2	57.8	42889	13	U10438	U10438 Caenorhabdi
c 562	19.2	57.8	2226	13	AF158747	AF158747 Drosophil	c 635	19.2	57.8	42968	13	CEY37A1A	299270 Caenorhabdi
c 563	19.2	57.8	2321	6	RATBEM3	D45414 Rat mRNA fo	c 636	19.2	57.8	43525	13	U88311	U88311 Caenorhabdi
c 564	19.2	57.8	2420	6	AY079517S2	AY079517 Mus muscu	c 637	19.2	57.8	43525	13	AF098501	AF098501 Caenorhab
c 565	19.2	57.8	2496	2	AK447985	AK447985 Sequence	c 638	19.2	57.8	44437	13	AC006645	AC006645 Caenorhab
c 566	19.2	57.8	2952	6	RNICA105	X92563 R. norvegicu	c 639	19.2	57.8	44614	13	CEC29E6	272504 Caenorhabdi
c 567	19.2	57.8	3358	4	AK065594	AK065594 Oryza sat	c 640	19.2	57.8	51235	13	AC006769	AC006769 Caenorhab
c 568	19.2	57.8	3397	6	RNU40652	U40652 Rattus norv	c 641	19.2	57.8	55613	5	AL512443	AL512443 Human DNA
c 569	19.2	57.8	3477	6	RATPDPFLP	D38222 Rattus sp.	c 642	19.2	57.8	55181	12	AC087443	AC087443 Homo sapi
c 570	19.2	57.8	4090	2	CQ588980	CQ588980 Sequence	c 643	19.2	57.8	65181	12	AC087443	AC087443 Homo sapi
c 571	19.2	57.8	4691	11	AY650211	AY650211 Seiurus a	c 644	19.2	57.8	67943	13	AC006722	AC006722 Caenorhab
c 572	19.2	57.8	4694	11	AY650193	AY650193 Dendroica	c 645	19.2	57.8	68270	13	CEY7A5A	AL034489 Caenorhab
c 573	19.2	57.8	8020	13	AC024832	AC024832 Caenorhab	c 646	19.2	57.8	69490	5	AC114294	AC114294 Homo sapi
c 574	19.2	57.8	8625	15	AF486545	AF486545 Campyloba	c 647	19.2	57.8	76022	12	AC168920	AC168920 Bos tauru
c 575	19.2	57.8	10048	13	CET26H8	282057 Caenorhabdi	c 648	19.2	57.8	83823	12	AC006789	AC006789 Caenorhab
c 576	19.2	57.8	10186	15	AF486550	AF486550 Campyloba	c 649	19.2	57.8	83943	2	CQ870148	CQ870148 Sequence
c 577	19.2	57.8	10187	15	AF486544	AF486544 Campyloba	c 650	19.2	57.8	86290	5	AC112503	AC112503 Homo sapi
c 578	19.2	57.8	10187	15	AF486551	AF486551 Campyloba	c 651	19.2	57.8	87428	13	CEY64G10A	AL110498 Caenorhab
c 579	19.2	57.8	10187	15	AF486558	AF486558 Campyloba	c 652	19.2	57.8	91927	5	AC004771	AC004771 Homo sapi
c 580	19.2	57.8	10627	2	I13706	I13706 Sequence 12	c 653	19.2	57.8	92817	5	AL591202	AL591202 Human DNA
c 581	19.2	57.8	12222	2	CS098060	CS098060 Sequence	c 654	19.2	57.8	94486	11	AC149069	AC149069 X. tropic
c 582	19.2	57.8	12222	2	AK390794	AK390794 Sequence	c 655	19.2	57.8	94762	12	AC157691	AC157691 Xenopus t
c 583	19.2	57.8	12222	5	HUMAIATP	K02212 Human alpha	c 656	19.2	57.8	95069	4	AP006265	AP006265 Oryza sat
c 584	19.2	57.8	14661	5	AY795074	AY795074 Homo sapi	c 657	19.2	57.8	97675	12	AC148534	AC148534 Macaca mu
c 585	19.2	57.8	15064	13	AC006713	AC006713 Caenorhab	c 658	19.2	57.8	98348	12	AC158848	AC158848 Bos tauru
c 586	19.2	57.8	15093	13	AF003149	AF003149 Caenorhab	c 659	19.2	57.8	99083	13	CEY37A1B	AL023835 Caenorhab
c 587	19.2	57.8	15106	6	AF521697	AF521697 Mus muscu	c 660	19.2	57.8	100883	12	AC157686	AC157686 Xenopus t
c 588	19.2	57.8	18572	13	CEY47G9	274035 Caenorhabdi	c 661	19.2	57.8	104970	13	AC024877	AC024877 Caenorhab
c 589	19.2	57.8	18806	13	CEY45F10C	299273 Caenorhabdi	c 662	19.2	57.8	110000	4	AP008213_070	Continuation (71 o
c 590	19.2	57.8	22528	13	CEZK1225	AL022289 Caenorhab	c 663	19.2	57.8	110000	4	AP008214_274	Continuation (275
c 591	19.2	57.8	24171	13	CEC31A11	283218 Caenorhabdi	c 664	19.2	57.8	110000	4	AP008217_210	Continuation (211
c 592	19.2	57.8	24330	13	AC024800	AC024800 Caenorhab	c 665	19.2	57.8	110000	4	AP008217_219	Continuation (220
c 593	19.2	57.8	24919	13	AF098991	AF098991 Caenorhab	c 666	19.2	57.8	110000	4	AP008218_085	Continuation (66 o
c 594	19.2	57.8	24950	13	AF014939	AF014939 Caenorhab	c 667	19.2	57.8	110000	4	CR382121_00	CR382121 Kluyverom
c 595	19.2	57.8	25406	13	U42437	U42437 Caenorhabdi	c 668	19.2	57.8	110000	4	AE016815_7	Continuation (8 of
c 596	19.2	57.8	25595	13	AF039045	AF039045 Caenorhab	c 669	19.2	57.8	110000	4	AP008207_252	Continuation (253
c 597	19.2	57.8	28586	13	CEC1A46	281470 Caenorhabdi	c 670	19.2	57.8	110000	4	AP008207_370	Continuation (371
c 598	19.2	57.8	29096	13	CEC05C12	268333 Caenorhabdi	c 671	19.2	57.8	110000	4	AP008211_057	Continuation (58 o
c 599	19.2	57.8	29344	13	CEY116A8B	AL021469 Caenorhab	c 672	19.2	57.8	110000	12	AC150790_0	AC150790 Bos tauru
c 600	19.2	57.8	29655	13	U40061	U40061 Caenorhabdi	c 673	19.2	57.8	110000	12	AC150872_0	AC150872 Bos tauru
c 601	19.2	57.8	29977	13	AF025461	AF025461 Caenorhab	c 674	19.2	57.8	110000	12	CEY111B2_1	Continuation (2 of
c 602	19.2	57.8	30767	13	AC006631	AC006631 Caenorhab	c 675	19.2	57.8	110000	15	CP000082_12	Continuation (13 o

676	19.2	57.8	110000	15	CP000082_13	Continuation (14 o	749	19.2	57.8	178434	12	AC118102
677	19.2	57.8	110000	15	CR931997_09	Continuation (10 o	750	19.2	57.8	180069	12	AC017041
678	19.2	57.8	110000	15	CR954246_25	Continuation (26 o	751	19.2	57.8	180742	5	AL359753
679	19.2	57.8	110000	15	BA000028_05	Continuation (6 of	752	19.2	57.8	181513	12	AC147749
680	19.2	57.8	117629	12	DQ369747_	DQ369747 Brassica	753	19.2	57.8	181914	12	AC024640
681	19.2	57.8	117919	4	AP003220	AP003220 Oryza sat	754	19.2	57.8	182356	6	AC154793
682	19.2	57.8	120984	5	AC006463	AC006463 Homo sapi	755	19.2	57.8	183046	2	CS086327
683	19.2	57.8	122855	13	AC084159	AC084159 Caenorhab	756	19.2	57.8	183099	12	AC034144
684	19.2	57.8	123496	12	AC178821	AC178821 Strongylo	757	19.2	57.8	183870	5	AC115284
685	19.2	57.8	123589	12	AC158852	AC158852 Bos tauru	758	19.2	57.8	184835	12	AC068091
686	19.2	57.8	124592	4	AC114828	AC114828 Oryza sat	759	19.2	57.8	186487	12	AC119833
687	19.2	57.8	129169	12	AC032038	AC032038 Homo sapi	760	19.2	57.8	186607	6	AL845292
688	19.2	57.8	131387	6	AC154403	AC154403 Mus muscu	761	19.2	57.8	187825	12	AC161061
689	19.2	57.8	134222	5	AB045363	AB045363 Homo sapi	762	19.2	57.8	187832	6	AC161596
690	19.2	57.8	137481	4	AP004704	AP004704 Oryza sat	763	19.2	57.8	188294	6	AC153608
691	19.2	57.8	138508	4	CNS08CCU	AL954156 Oryza sat	764	19.2	57.8	189610	5	AC118584
692	19.2	57.8	140620	4	AP003415	AP003415 Oryza sat	765	19.2	57.8	190628	12	AC178683
693	19.2	57.8	140788	12	AC148419	AC148419 Dasyypus n	766	19.2	57.8	190673	6	AL645808
694	19.2	57.8	141027	4	AC132486	AC132486 Oryza sat	767	19.2	57.8	193188	12	AC006884
695	19.2	57.8	142938	6	AC118735	AC118735 Mus muscu	768	19.2	57.8	193301	5	CNS01DRJ
696	19.2	57.8	145514	4	AP003681	AP003681 Oryza sat	769	19.2	57.8	193909	6	AC115797
697	19.2	57.8	145697	5	AC020642	AC020642 Homo sapi	770	19.2	57.8	196383	12	AC113159
698	19.2	57.8	146570	12	AC149792	AC149792 Aedes aeg	771	19.2	57.8	196947	6	AC121843
699	19.2	57.8	146927	12	AC020144	AC020144 Drosophil	772	19.2	57.8	198349	6	AC151298
700	19.2	57.8	147423	12	AC177333	AC177333 Strongylo	773	19.2	57.8	199239	6	AC131718
701	19.2	57.8	147505	5	CNS01D7A	AL132708 Human chr	774	19.2	57.8	199631	12	AC175342
702	19.2	57.8	148120	12	AC141805	AC141805 Apis mell	775	19.2	57.8	200898	4	AC092058
703	19.2	57.8	150789	12	AC006685	AC006685 Caenorhab	776	19.2	57.8	200968	5	AP005516
704	19.2	57.8	152272	4	AC137589	AC137589 Oryza sat	777	19.2	57.8	201031	6	AC154356
705	19.2	57.8	152272	4	AC144558	AC144558 Oryza sat	778	19.2	57.8	202719	12	AC111339
706	19.2	57.8	152637	12	AC006734	AC006734 Caenorhab	779	19.2	57.8	203950	12	AC127649
707	19.2	57.8	153388	4	AP003289	AP003289 Oryza sat	780	19.2	57.8	209480	12	AC140965
708	19.2	57.8	153587	12	AC148751	AC148751 Rhinolo	781	19.2	57.8	209583	6	AC165442
709	19.2	57.8	154482	6	AC126430	AC126430 Mus muscu	782	19.2	57.8	210196	5	AC016717
710	19.2	57.8	156069	4	AC136842	AC136842 Oryza sat	783	19.2	57.8	210558	5	AC016717
711	19.2	57.8	156312	5	AC138470	AC138470 Homo sapi	784	19.2	57.8	211910	6	AC116573
712	19.2	57.8	156320	2	CS162220	CS162220 Sequence	785	19.2	57.8	212071	6	AC136754
713	19.2	57.8	156331	5	AC074331	AC074331 Homo sapi	786	19.2	57.8	212309	6	AC154628
714	19.2	57.8	156342	5	AC158566	AC158566 Mus muscu	787	19.2	57.8	212682	12	AC157414
715	19.2	57.8	156958	5	AC013460	AC013460 Homo sapi	788	19.2	57.8	213401	12	AC124940
716	19.2	57.8	157722	13	AC092248	AC092248 Drosophil	789	19.2	57.8	213473	6	AC140287
717	19.2	57.8	159420	6	CT025589	CT025589 Mouse DNA	790	19.2	57.8	215664	12	AC129367
718	19.2	57.8	160163	12	AC180897	AC180897 Strongylo	791	19.2	57.8	216132	12	AC146533
719	19.2	57.8	160736	11	BX530031	BX530031 Zebrafish	792	19.2	57.8	216935	12	AC147486
720	19.2	57.8	161054	5	AC104444	AC104444 Homo sapi	793	19.2	57.8	218381	12	AC150564
721	19.2	57.8	162397	12	AC012536	AC012536 Homo sapi	794	19.2	57.8	222520	12	AC120336
722	19.2	57.8	162625	12	CET08E3	Z83124 Caenorhabdi	795	19.2	57.8	222895	12	AC084065
723	19.2	57.8	163396	5	AL161445	AL161445 Human DNA	796	19.2	57.8	222962	12	AC094146
724	19.2	57.8	163528	6	AC121951	AC121951 Mus muscu	797	19.2	57.8	223110	6	AC123752
725	19.2	57.8	165871	5	CNS01DVO	AL135998 Human chr	798	19.2	57.8	224009	12	AC131058
726	19.2	57.8	167336	12	AC153669	AC153669 Bos tauru	799	19.2	57.8	224318	12	AC160377
727	19.2	57.8	167475	6	AC090881	AC090881 Mus Muscu	800	19.2	57.8	227606	13	CEY53F4B
728	19.2	57.8	167576	12	AC147297	AC147297 Pan trogl	801	19.2	57.8	227607	6	AC120570
729	19.2	57.8	167724	12	AC147577	AC147577 Orolemur	802	19.2	57.8	227940	12	AC151108
730	19.2	57.8	168821	6	AC138638	AC138638 Mus muscu	803	19.2	57.8	228943	12	AC152222
731	19.2	57.8	168946	11	CR847948	CR847948 Zebrafish	804	19.2	57.8	229867	12	AC091287
732	19.2	57.8	169226	12	AC006913	AC006913 Caenorhab	805	19.2	57.8	230058	12	AC108229
733	19.2	57.8	169246	5	AC091815	AC091815 Homo sapi	806	19.2	57.8	230879	5	AP003465
734	19.2	57.8	170094	5	AC146033	AC146033 Pan trogl	807	19.2	57.8	231778	11	BX511175
735	19.2	57.8	170164	12	CT573031	CT573031 Mus muscu	808	19.2	57.8	232415	12	AC099429
736	19.2	57.8	170993	12	AC021528	AC021528 Homo sapi	809	19.2	57.8	232571	6	AC116557
737	19.2	57.8	171141	6	AC157651	AC157651 Mus muscu	810	19.2	57.8	232830	12	CT033781
738	19.2	57.8	171787	5	AC100777	AC100777 Homo sapi	811	19.2	57.8	235045	12	AC119588
739	19.2	57.8	172942	12	AC016899	AC016899 Homo sapi	812	19.2	57.8	235071	12	AC115654
740	19.2	57.8	172959	12	AC119422	AC119422 Papio anu	813	19.2	57.8	235720	12	AC120712
741	19.2	57.8	173556	12	AC074388	AC074388 Homo sapi	814	19.2	57.8	235852	6	AC167244
742	19.2	57.8	173803	12	CNS01DV2	AL133448 Homo sapi	815	19.2	57.8	236144	13	AE003643
743	19.2	57.8	173992	12	CT030660	CT030660 Mus muscu	816	19.2	57.8	237374	6	AC112030
744	19.2	57.8	174298	12	AC167299	AC167299 Oryctolag	817	19.2	57.8	238722	12	AC098108
745	19.2	57.8	174642	5	AC096751	AC096751 Homo sapi	818	19.2	57.8	239591	12	AC128432
746	19.2	57.8	174790	12	AC181037	AC181037 Strongylo	819	19.2	57.8	240309	12	AC103118
747	19.2	57.8	175396	12	AC156621	AC156621 Mus muscu	820	19.2	57.8	240931	12	AC107097
748	19.2	57.8	176584	12	AC027812	AC027812 Homo sapi	821	19.2	57.8	245805	12	AC006752

C 822	19.2	57.8	248510	12	AC095814	AC095814 Rattus no	C 895	19	57.2	110000	15	AE017285_20	Continuation (21 o
C 823	19.2	57.8	249200	12	AC094799	AC094799 Rattus no	C 896	19	57.2	111973	12	AP007350	AP007350 Lotus cor
C 824	19.2	57.8	249262	12	AC079430	Mus muscu	C 897	19	57.2	117217	5	AP002076	AP002076 Homo sapi
C 825	19.2	57.8	249481	12	AC159672	Bos tauru	C 898	19	57.2	117752	5	AC004743	AC004743 Homo sapi
C 826	19.2	57.8	251742	12	AC096526	Rattus no	C 899	19	57.2	125860	4	AP004151	AP004151 Oryza sat
C 827	19.2	57.8	252474	12	AC093999	Rattus no	C 900	19	57.2	127006	5	AL138849	AL138849 Human DNA
C 828	19.2	57.8	254497	12	AC119520	Rattus no	C 901	19	57.2	129690	6	AC138601	AC138601 Mus muscu
C 829	19.2	57.8	255305	12	AC160576	Bos tauru	C 902	19	57.2	133878	12	AC136064	AC136064 Rattus no
C 830	19.2	57.8	256719	6	AC154873	Mus muscu	C 903	19	57.2	140056	12	BS511131	BS511131 Danio rer
C 831	19.2	57.8	257158	12	AC168001	Bos tauru	C 904	19	57.2	145770	12	AC148384	AC148384 Dasyypus n
C 832	19.2	57.8	257703	13	CEX111B2A	AL132904 Caenorhab	C 905	19	57.2	146466	5	AL353637	AL353637 Human DNA
C 833	19.2	57.8	257728	12	AC006846	Caenorhab	C 906	19	57.2	147505	6	BS572085	BS572085 Mouse DNA
C 834	19.2	57.8	258174	12	AC079429	Mus muscu	C 907	19	57.2	148347	4	AP005408	AP005408 Oryza sat
C 835	19.2	57.8	258274	12	AC105595	Rattus no	C 908	19	57.2	148683	11	BS248499	BS248499 Zebrafish
C 836	19.2	57.8	259415	12	AC156178	Bos tauru	C 909	19	57.2	148821	6	CT025643	CT025643 Mouse DNA
C 837	19.2	57.8	259679	12	AC169684	Bos tauru	C 910	19	57.2	148866	12	AC176538	AC176538 Strongylo
C 838	19.2	57.8	259701	12	AC164709	Bos tauru	C 911	19	57.2	149103	12	AC128053	AC128053 Rattus no
C 839	19.2	57.8	259824	12	AC098374	Rattus no	C 912	19	57.2	150399	5	AC093680	AC093680 Homo sapi
C 840	19.2	57.8	261771	12	AC125864	Rattus no	C 913	19	57.2	151041	4	AP003258	AP003258 Oryza sat
C 841	19.2	57.8	263335	15	CJ11168X5	AL139078 Campyloba	C 914	19	57.2	151345	4	AC135919	AC135919 Oryza sat
C 842	19.2	57.8	279031	12	AC176560	Strongylo	C 915	19	57.2	153006	12	AC027782	AC027782 Homo sapi
C 843	19.2	57.8	279504	12	AC169888	Bos tauru	C 916	19	57.2	155021	12	AC103385	AC103385 Mus muscu
C 844	19.2	57.8	282690	12	AC164710	Bos tauru	C 917	19	57.2	159301	12	CT027752	CT027752 Danio rer
C 845	19.2	57.8	287648	12	AC134140	Rattus no	C 918	19	57.2	160599	11	AL954363	AL954363 Zebrafish
C 846	19.2	57.8	289905	12	AC094576	Rattus no	C 919	19	57.2	162415	6	AC153986	AC153986 Mus muscu
C 847	19.2	57.8	292390	12	AC105677	Rattus no	C 920	19	57.2	162895	12	AC178016	AC178016 Strongylo
C 848	19.2	57.8	293024	12	CEY5988	Z98870 Caenorhabdi	C 921	19	57.2	163186	5	AC015994	AC015994 Homo sapi
C 849	19.2	57.8	299782	12	AC006844	Caenorhab	C 922	19	57.2	163213	12	AC016984	AC016984 Mus muscu
C 850	19.2	57.8	300000	5	AP002534	AP002534 Homo sapi	C 923	19	57.2	164137	12	AC129854	AC129854 Rattus no
C 851	19.2	57.8	322101	2	AX814520	AX814520 Sequence	C 924	19	57.2	165328	6	AC121782	AC121782 Mus muscu
C 852	19.2	57.8	330724	12	CEY67H2	AC022475 Caenorhab	C 925	19	57.2	165799	5	AL358789	AL358789 Human DNA
C 853	19.2	57.8	336873	12	AC073666	AC073666 Mus muscu	C 926	19	57.2	167454	12	AC021306	AC021306 Homo sapi
C 854	19.2	57.8	348986	15	BS572093	BS572093 Prochloro	C 927	19	57.2	167546	12	AC163985	AC163985 Rhinolph
C 855	19.2	57.8	349980	2	Q870293	Q870293 Sequence	C 928	19	57.2	167662	5	AC055753	AC055753 Homo sapi
C 856	19	57.2	95	15	STWATROA	M14388 S.lividans	C 929	19	57.2	169147	12	AC129989	AC129989 Rattus no
C 857	19	57.2	184	2	AX905690	AX905690 Sequence	C 930	19	57.2	169547	13	CP000189	CP000189 Drosophil
C 858	19	57.2	184	2	BD041223	BD041223 Sequence	C 931	19	57.2	169742	6	AL646101	AL646101 Mouse DNA
C 859	19	57.2	184	2	AR741484	AR741484 Sequence	C 932	19	57.2	169928	5	AC019067	AC019067 Homo sapi
C 860	19	57.2	386	7	HSB3311XE9	Z67628 H.sapiens D	C 933	19	57.2	170072	6	AC0116329	AC0116329 Mus muscu
C 861	19	57.2	580	7	BV043651	BV043651 S12P6568	C 934	19	57.2	172697	12	AC076965	AC076965 Homo sapi
C 862	19	57.2	1531	13	AY278950	AY278950 Branchios	C 935	19	57.2	173446	12	AC119178	AC119178 Mus muscu
C 863	19	57.2	2413	5	AK095951	AK095951 Homo sapi	C 936	19	57.2	173474	6	AC127686	AC127686 Mus muscu
C 864	19	57.2	2771	11	CR523878	CR523878 Gallus ga	C 937	19	57.2	174714	12	AC111112	AC111112 Mus muscu
C 865	19	57.2	3325	5	BC009208	BC009208 Homo sapi	C 938	19	57.2	175560	12	AC171304	AC171304 Bos tauru
C 866	19	57.2	3355	5	BC058496	BC058496 Homo sapi	C 939	19	57.2	175920	12	AC164351	AC164351 Bos tauru
C 867	19	57.2	3915	2	AX250162	AX250162 Sequence	C 940	19	57.2	177255	12	AC123235	AC123235 Rattus no
C 868	19	57.2	4136	2	AX455340	AX455340 Sequence	C 941	19	57.2	177800	6	AC157095	AC157095 Mus muscu
C 869	19	57.2	4151	2	CQ492548	CQ492548 Sequence	C 942	19	57.2	177897	11	AC149879	AC149879 Xenopus t
C 870	19	57.2	4151	2	CQ492548	CQ492548 Sequence	C 943	19	57.2	178013	4	AC136219	AC136219 Oryza sat
C 871	19	57.2	6452	15	AY345225	AY345225 Arsenite-	C 944	19	57.2	180227	6	AC138358	AC138358 Mus muscu
C 872	19	57.2	12275	13	AF063309	AF063309 Caenorhab	C 945	19	57.2	180742	6	CT572998	CT572998 Mouse DNA
C 873	19	57.2	16120	11	AP004807	AP004807 Elops sau	C 946	19	57.2	180892	6	AC125533	AC125533 Mus muscu
C 874	19	57.2	16713	11	AB051070	AB051070 Elops haw	C 947	19	57.2	181127	6	AL672038	AL672038 Mouse DNA
C 875	19	57.2	42551	12	CT005259_7	Continuation (8 of	C 948	19	57.2	181544	12	AC167908	AC167908 Bos tauru
C 876	19	57.2	45800	11	AC147865	AC147865 Gopherus	C 949	19	57.2	182189	12	AC118424	AC118424 Rattus no
C 877	19	57.2	48768	5	AL157933	AL157933 Human DNA	C 950	19	57.2	182579	12	AC167513	AC167513 Oryctolag
C 878	19	57.2	56013	6	AL773581	AL773581 Mouse DNA	C 951	19	57.2	182627	5	AC096725	AC096725 Homo sapi
C 879	19	57.2	66944	4	MIATGENA_3	Continuation (4 of	C 952	19	57.2	183094	6	AC087903	AC087903 Mus muscu
C 880	19	57.2	73986	12	CT030037_3	Continuation (4 of	C 953	19	57.2	185541	6	AL844537	AL844537 Mouse DNA
C 881	19	57.2	91246	12	LMFLCHR18_10	Continuation (11 o	C 954	19	57.2	185786	12	AC079372	AC079372 Mus muscu
C 882	19	57.2	99878	12	AC087108	AC087108 Homo sapi	C 955	19	57.2	186585	12	AC161803	AC161803 Mus muscu
C 883	19	57.2	100074	6	AL732548	AL732548 Mouse DNA	C 956	19	57.2	190254	11	BS511223	BS511223 Zebrafish
C 884	19	57.2	108457	5	HSJ800321	AL109955 Human DNA	C 957	19	57.2	190469	6	AC109172	AC109172 Mus muscu
C 885	19	57.2	110000	4	AP008207_139	Continuation (140	C 958	19	57.2	191685	6	AC131663	AC131663 Mus muscu
C 886	19	57.2	110000	4	AP008208_142	Continuation (143	C 959	19	57.2	193164	12	AC179142	AC179142 Strongylo
C 887	19	57.2	110000	4	AP008208_143	Continuation (144	C 960	19	57.2	193304	12	AC109370	AC109370 Rattus no
C 888	19	57.2	110000	4	AP008211_118	Continuation (119	C 961	19	57.2	193386	6	AC116738	AC116738 Mus muscu
C 889	19	57.2	110000	4	AP008211_119	Continuation (120	C 962	19	57.2	194437	12	AC117948	AC117948 Homo sapi
C 890	19	57.2	110000	4	AP008211_185	Continuation (186	C 963	19	57.2	196798	6	AC163035	AC163035 Mus muscu
C 891	19	57.2	110000	12	CT005430_0	AC099430 Rattus no	C 964	19	57.2	199287	6	AC125043	AC125043 Mus muscu
C 892	19	57.2	110000	12	CT005259_6	Continuation (7 of	C 965	19	57.2	199987	12	AC171512	AC171512 Bos tauru
C 893	19	57.2	110000	12	LMFLCHR18_09	Continuation (10 o	C 966	19	57.2	201304	6	AL929012	AL929012 Mouse DNA
C 894	19	57.2	110000	15	AE000513_08	Continuation (9 of	C 967	19	57.2	201748	12	AC161914	AC161914 Bos tauru

968	19	57.2	202023	12	AC141270	1
C 969	19	57.2	202333	12	AC165202	
C 970	19	57.2	203604	6	AC158963	Mus muscu
971	19	57.2	204206	12	AC176020	Strongylo
972	19	57.2	205101	6	AL645564	Mouse DNA
973	19	57.2	206187	12	AC021172	Homo sapi
C 974	19	57.2	207223	6	AL732521	Mouse DNA
C 975	19	57.2	207450	12	AC133253	Rattus no
C 976	19	57.2	207782	4	AP003764	Oryza sat
C 977	19	57.2	208140	6	AC131084	Mus muscu
C 978	19	57.2	209409	12	AC164361	Bos tauru
C 979	19	57.2	209885	6	AC153512	Mus muscu
C 980	19	57.2	210190	6	AC115781	Mus muscu
981	19	57.2	211056	6	AL591207	Mouse DNA
C 982	19	57.2	211061	6	AC153591	Mus muscu
983	19	57.2	211190	12	AC152229	Bos tauru
984	19	57.2	211776	12	AC123347	Rattus no
C 985	19	57.2	212081	11	BX005083	Zebrafish
C 986	19	57.2	213783	12	AC098119	Rattus no
987	19	57.2	215113	6	AC105167	Mus muscu
C 988	19	57.2	216294	6	AC154470	Mus muscu
C 989	19	57.2	220632	6	AC136517	Mus muscu
990	19	57.2	221254	6	AC135258	Mus muscu
991	19	57.2	223432	5	AC126755	Homo sapi
992	19	57.2	227026	6	AC113476	Mus muscu
993	19	57.2	227946	6	AC124175	Mus muscu
C 994	19	57.2	230156	12	AC132773	Rattus no
C 995	19	57.2	234584	12	AC111604	Rattus no
C 996	19	57.2	235763	12	AC096378	Rattus no
997	19	57.2	238247	12	AC095452	Rattus no
998	19	57.2	240349	12	AC096133	Rattus no
C 999	19	57.2	242228	12	AC135821	Rattus no
C1000	19	57.2	243676	6	AC115763	Mus muscu

ALIGNMENTS

```

RESULT 1
BD233463
LOCUS       603 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION  Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION   BD233463
VERSION     BD233463.1  GI:33043233
KEYWORDS    JP 2002519016-A/9.
SOURCE      Homo sapiens (human)
ORIGIN      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1  (bases 1 to 603)
AUTHORS    Kato,S. and Kimura,T.
TITLE      Human protein having hydrophobic domain and DNA encoding the same
JOURNAL    Patent: JP 2002519016-A 9 02-JUL-2002;
COMMENT    SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
            OS   Homo sapiens (human)
            PN   JP 2002519016-A/9
            PD   02-JUL-2002
            PF   18-JUN-1999  JP 2000557267
            PI   SEISHI KATO,TOMOKO KIMURA
            PC   C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N5/00,C12N5/
            00
            CC   Human protein having hydrophobic domain and DNA encoding the
            CC   same
            FH   Key      Location/Qualifiers
            FT   source   1..603
            FT              /organism='Homo sapiens (human)' .
            FT              Location/Qualifiers
            FT                1..603
            FT                  /organism='Homo sapiens'
            FT                  /mol_type='genomic DNA'
            FT                  /db_xref='taxon:9606'

FEATURES             source

```

```

ORIGIN
Query Match          90.4%; Score 30; DB 2; Length 603;
Best Local Similarity 64.3%; Pred. No. 0.0065;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
   |||:||| |||:|||:|||:|||:|||:|||
Db 550 GCCCAGCGTGGCCGCGCCCTTACAAAGACTATGAAGTAAAG 591

RESULT 2
BD209699          1212 bp DNA linear PAT 17-JUL-2003
LOCUS             Compositions isolated from skin cells and methods for their use.
DEFINITION
ACCESSION         BD209699
VERSION           BD209699.1 GI:33019469
KEYWORDS          JP 2002512798-A/171.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                  Homiidae; Homo.
REFERENCE
1 (bases 1 to 1212)
Strachan,L., Sleeman,M., Watson,J.D., Orrust,R., Kumble,A. and
Murison,J.G.
TITLE             Compositions isolated from skin cells and methods for their use
JOURNAL           Patent: JP 2002512798-A 171 08-MAY-2002;
                  GENESIS RESEARCH AND DEVELOPMENT CORP LTD
COMMENT           OS Homo sapiens (human)
                  PN JP 2002512798-A/171
                  PD 08-MAY-2002
                  PF 29-APR-1999 JP 2000546009
                  PR 29-APR-1998 US 09/069726, 09-NOV-1998 US 09/188930 PI
                  LORNA STRACHAN, MATTHEW SLEEMAN, JAMES DOUGLAS WATSON, RENE PI
                  ONRUST,
                  ANAND KUMBLE, JAMES GREG MURISON
                  PC C12N15/09, A61K38/00, A61P9/00, A61P17/00, A61P29/00, A61P31/18, PC
                  A61P35/00,
                  PC C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12N15/
                  00,
                  PC A61K37/02, C12N5/00
                  CC Compositions isolated from skin cells and methods for their
                  use.
                  CH Key Location/Qualifiers
                  FT source 1..1212
                  FT Location/Qualifiers
                  1..1212
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"

ORIGIN
Query Match          90.4%; Score 30; DB 2; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.0074;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
   |||:||| |||:|||:|||:|||:|||:|||
Db 990 GCCCAGCGTGGCCGCGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 3
AR341505          1212 bp DNA linear PAT 17-AUG-2003
LOCUS             Sequence 249 from patent US 6573095.
DEFINITION
ACCESSION         AR341505
VERSION           AR341505.1 GI:33733640
KEYWORDS          .
SOURCE            Unknown.
ORGANISM          Unknown.

```



```

Unclassified.
REFERENCE
1 (bases 1 to 1212)
AUTHORS
Strachan,L.
TITLE
Polynucleotides isolated from skin cells
JOURNAL
Patent: US 6573095-A 249 03-JUN-2003;
Genesis Research & Development Corporation Limited; Parnell;
NZX;
FEATURES
source
Location/Qualifiers
1..1212
/mol_type="genomic DNA"
ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.0074;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCGNTGGCCNTCCNTAYAARGAYTAGTGCTNAA 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 990 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 4
AX078375 1619 bp DNA linear PAT 22-FEB-2001
LOCUS
DEFINITION
Sequence 43 from Patent WO0107612.
ACCESSION
AX078375
VERSION
AX078375.1 GI:13158044
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
LOCATION/Qualifiers
1..1619
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incye ID No: 2681738CB1"
ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 1619;
Best Local Similarity 64.3%; Pred. No. 0.0079;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCGNTGGCCNTCCNTAYAARGAYTAGTGCTNAA 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1137 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1178

RESULT 5
BD233473 1718 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION
Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION
BD233473
VERSION
BD233473.1 GI:33043243
KEYWORDS
JP 2002519016-A/19.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
LOCATION/Qualifiers
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1718)
AUTHORS
Kato,S. and Kimura,T.
TITLE
Human protein having hydrophobic domain and DNA encoding the same
JOURNAL
Patent: JP 2002519016-A 19 02-JUL-2002;

```

```

SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
OS Homo sapiens (human)
PN JP 2002519016-A/19
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO,TOMOKO KIMURA
PC
C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N5/00,C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
CC same
FH Key 1..1718
FT source Location/Qualifiers
FT /organism="Homo sapiens (human)".
FEATURES
source
Location/Qualifiers
1..1718
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 1718;
Best Local Similarity 64.3%; Pred. No. 0.0079;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCGNTGGCCNTCCNTAYAARGAYTAGTGCTNAA 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 561 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 602

RESULT 6
AK172760 1826 bp mRNA linear PRI 07-MAY-2004
LOCUS
DEFINITION
Homo sapiens cDNA FLJ23921 fis, clone COL02043, highly similar to
Homo sapiens retinoic acid induced 3 (RAI3).
ACCESSION
AK172760
VERSION
AK172760.1 GI:47077732
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
LOCATION/Qualifiers
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obyaashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE
NEDO human cDNA sequencing project
REFERENCE
2 (bases 1 to 1826)
JOURNAL
Sugano,S. and Suzuki,Y.
Direct Submission
Submitted (22-APR-2004) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
[E-mail:flcdna@mail.ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416]
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction,
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partially supported by Science and Technology Agency).
FEATURES
source
Location/Qualifiers
1..1826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL02043"
/tissue_type="colon"
/clone_lib="COL"
/notes="cloning vector pME18SFL3"
ORIGIN

```



```

Query Match          90.4%; Score 30; DB 5; Length 1826;
Best Local Similarity 64.3%; Pred. No. 0.008;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYARGAYTAYGAGTNAAR 42
Db 638 GCCACGCTTGGCCGAGCCCTTACAAGACTATGTAAGTAAAG 679

RESULT 7
LOCUS CQ723177
DEFINITION Sequence 9111 from Patent WO02068579.
ACCESSION CQ723177
VERSION CQ723177.1 GI:42284034
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 9111 06-SEP-2002;
PE Corporation (NY) (US)
LOCATION/Qualifiers
1. .2290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source
1. .2290
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MG:923 IMAGE:2988011"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH MGC 15"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .2296
/genes="GPCR5A"
/notes="synonyms: RAIG1, GPCR5A, RA13"
/db_xref="GeneID:9052"
/db_xref="MIM:604138"
95. .1168
/genes="GPCR5A"
/codon_start=1
/product="G protein-coupled receptor, family C, group 5,
member A"
/protein_id="AAH03665.1"
/db_xref="GI:13177796"
/db_xref="GeneID:9052"
/db_xref="MIM:604138"
/translation="MATTVPDGCGRGLSKYKYLCDKAEAWGIVLETATAGVVTVA
FMLTPLIVCKQDSNRRKMLPTQFLFGLVGLTAFIIGLDGSTGPTFFLFG
ILFSICGLAHAVSLKLVGRKPLSLVILGLAVGSLVDVIAEIVLTMNT
NNVSELSAPRNEFDVLLTYVLFMLALTFLMSSFTFCGSGTGMKRGHAHLYTML
LSIAIWWAITLMLPFDPRDDTILSSALAANGWVFLLAYVSPFWLLTKGRNPD
YFVEDAFCKPLQVKSGYGVENRAYSQEITQGEETGDTLYAPYSTHFLQNPQPKKE
FSIPRAHAWPSPYKDYEVKKEGS"

ORIGIN
Query Match          90.4%; Score 30; DB 5; Length 2296;
Best Local Similarity 64.3%; Pred. No. 0.0084;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYARGAYTAYGAGTNAAR 42
Db 1112 GCCACGCTTGGCCGAGCCCTTACAAGACTATGTAAGTAAAG 1153

TITLE
JOURNAL PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:13177795.
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: anadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettaman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 3 Row: 1 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 12056470.
Location/Qualifiers
1. .2296
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MG:923 IMAGE:2988011"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH MGC 15"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .2296
/genes="GPCR5A"
/notes="synonyms: RAIG1, GPCR5A, RA13"
/db_xref="GeneID:9052"
/db_xref="MIM:604138"
95. .1168
/genes="GPCR5A"
/codon_start=1
/product="G protein-coupled receptor, family C, group 5,
member A"
/protein_id="AAH03665.1"
/db_xref="GI:13177796"
/db_xref="GeneID:9052"
/db_xref="MIM:604138"
/translation="MATTVPDGCGRGLSKYKYLCDKAEAWGIVLETATAGVVTVA
FMLTPLIVCKQDSNRRKMLPTQFLFGLVGLTAFIIGLDGSTGPTFFLFG
ILFSICGLAHAVSLKLVGRKPLSLVILGLAVGSLVDVIAEIVLTMNT
NNVSELSAPRNEFDVLLTYVLFMLALTFLMSSFTFCGSGTGMKRGHAHLYTML
LSIAIWWAITLMLPFDPRDDTILSSALAANGWVFLLAYVSPFWLLTKGRNPD
YFVEDAFCKPLQVKSGYGVENRAYSQEITQGEETGDTLYAPYSTHFLQNPQPKKE
FSIPRAHAWPSPYKDYEVKKEGS"

ORIGIN
Query Match          90.4%; Score 30; DB 5; Length 2296;
Best Local Similarity 64.3%; Pred. No. 0.0084;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYARGAYTAYGAGTNAAR 42
Db 1112 GCCACGCTTGGCCGAGCCCTTACAAGACTATGTAAGTAAAG 1153

```

```
RESULT 9
LOCUS AF506289 2297 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens orphan G protein-coupling receptor PEIG-1 mRNA,
complete cds.
ACCESSION AF506289
VERSION AF506289.1 GI:21779962
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2040)
Cafferata, E.G., Gonzalez-Guerrico, A.M., Pivetta, O.H. and
Santa-Coloma, T.A.
Identification by differential display of a mRNA specifically
induced by 12-O-tetradecanoylphorbol-13-acetate (TPA) in T84 human
colon carcinoma cells
Cell. Mol. Biol. 42 (5), 797-804 (1996)
JOURNAL
PUBMED 8832110
REFERENCE
2 (bases 1 to 2297)
Cafferata, E.G., Gonzalez-Guerrico, A.M., Costanzo, R., Pivetta, O.H.
and Santa-Coloma, T.A.
Direct Submission
Submitted (24-APR-2002) Laboratorio de Biologia Celular y
Molecular, Instituto de Investigaciones Bioquimicas Fundacion
Campomar, Patricia Argentina 435, Buenos Aires 1405, Argentina
Location/Qualifiers
1. .2297
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12.3"
/sex="male"
/cell_line="T84"
/tissue_type="colorectal carcinoma"
/notes="derived from dbEST BE519991"
108..1181
/notes="belongs to the type 3 G protein-coupling receptor
family; characterized by a 7-transmembrane domain motif;
phorbol ester induced protein-1"
/codon_start=1
/product="orphan G protein-coupling receptor PEIG-1"
/protein_id="AA077594.1"
/db_xref="GI:21779963"
/translacion="MATTVPDGCNGLSKYRLCDKAAWGIIVLETATAGVVTVA
FMULTPLVCKVQDSNRKMLPTQFLVLGLVGLTFAFIIGLDGSGTPTREFLFG
ILFSICSLAHAVSLTKLVGRKPLSLVLGLAVGFLSLVQDVAIEYIVLPMNR
NNVVFSELSAPRRNEDFVLLLTLYVFLMALTFMLSSFTFCGSGTGWKRHGAHYLTML
LSIAIWTATLMLPDRRWDITLSALAAAGWFLLYVSPEFWLTKQNPMD
YPVEDAFCKPLVKGYGVENRAYSQEITQGETGTLTYAPYSTHFLQNPQPOKE
FSIPRAHAFSPYKDYEVKKEGS"
ORIGIN
Query Match 90.4%; Score 30; DB 5; Length 2297;
Best Local Similarity 64.3%; Pred. No. 0.0084;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
|||:|||||
Db 1125 GCCACGCTGGCGGAGCCCTTACAAAGACTATGAGTAAAG 1166
|||||:|||||
RESULT 11
LOCUS AF095448 2302 bp mRNA linear PRI 29-DEC-1998
DEFINITION Homo sapiens putative G protein-coupled receptor (RAIG1) mRNA,
complete cds.
ACCESSION AF095448
VERSION AF095448.1 GI:4063889
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2302)
Cheng, Y. and Lotan, R.
Molecular cloning and characterization of a novel retinoic
acid-inducible gene that encodes a putative G protein-coupled
receptor
J. Biol. Chem. 273 (52), 35008-35015 (1998)
JOURNAL
PUBMED 9857033
REFERENCE
2 (bases 1 to 2302)
Cheng, Y. and Lotan, R.
Direct Submission
Submitted (27-SEP-1998) Tumor Biology, The University of Texas M.
D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX
77030, USA
Location/Qualifiers
1. .2302
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p12.3-p13"
1..2302
/genes="RAIG1"
/notes="induced by retinoic acid"
100..1173
/genes="RAIG1"
/codon_start=1
/product="putative G protein-coupled receptor"
/protein_id="AAC98506.1"
/db_xref="GI:4063890"
/translacion="MATTVPDGCNGLSKYRLCDKAAWGIIVLETATAGVVTVA
FMULTPLVCKVQDSNRKMLPTQFLVLGLVGLTFAFIIGLDGSGTPTREFLFG
ILFSICSLAHAVSLTKLVGRKPLSLVLGLAVGFLSLVQDVAIEYIVLPMNR
NNVVFSELSAPRRNEDFVLLLTLYVFLMALTFMLSSFTFCGSGTGWKRHGAHYLTML
NNVVFSELSAPRRNEDFVLLLTLYVFLMALTFMLSSFTFCGSGTGWKRHGAHYLTML"
ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0084;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
|||:|||||
Db 1117 GCCACGCTGGCGGAGCCCTTACAAAGACTATGAGTAAAG 1158
|||||:|||||
RESULT 10
LOCUS AX930411 2302 bp DNA linear PAT 22-DEC-2003
DEFINITION Sequence 2 from Patent WO03087832.
ACCESSION AX930411
VERSION AX930411.1 GI:40312209
```

LSIAIWAVITLLMLPDDRRWDDTILSSALAANGVFLLAYVSPFWLLTKORNPMD
YPVEDAFCKPQLVKKSYGVENRAYSQEETIQGFEETGDTLYAPYSTHFQLOQPPOKE
FSIPRAHAWPSPKDYEVKKEGS"

ORIGIN

Query Match 90.4%; Score 30; DB 5; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0084;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYARGAYTAYGARGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

Db 1117 GCCCAGCGCTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1158

RESULT 12

BD156680

LOCUS 2446 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156680

VERSION BD156680.1 GI:27862438

KEYWORDS JP 2002191363-A/11523

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominiidae; Homo.

1 (bases 1 to 2446)

Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,

Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 11523 09-JUL-2002;

HELIIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11523

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU

PI SAITO,

PI JIUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers

FT CDS (254)..(1324).

Location/Qualifiers

1..2446

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 90.4%; Score 30; DB 2; Length 2446;
Best Local Similarity 64.3%; Pred. No. 0.0085;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYARGAYTAYGARGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

Db 1271 GCCCAGCGCTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1312

RESULT 13

AX877483

LOCUS 2446 bp DNA linear PAT 17-DEC-2003

DEFINITION Sequence 12388 from Patent EP1074617.

ACCESSION AX877483

VERSION AX877483.1 GI:40032219

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE

AUTHORS

1

Ota, T.,

Ishii, S.,

Sugiyama, T.,

Wakamatsu, A.,

Nagai, K. and

Otsuki, T.

Primer for synthesizing full-length cDNA and their use

Patent: EP 1074617-A 12388 07-FEB-2001;

Research Association for Biotechnology (JP)

Location/Qualifiers

1..2446

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

254..1327

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE89746.1"

/db_xref="GI:40032220"

/translation="MATVDPGCRNGLKSKYVRLCDKAEANGIVLETVATAGVVTVA

FMLTPILVKVQDSNRRLMPTQFLILGLVIGLTFATFIIGLDGTGPTFFFLFG

ILFSICFSCILAHAVSLTKLVRKPLSLVILGLAVGFSLVQDVIAIYVLTNRT

NVNVESELAPRNEFDVLLTVLFMALTFMSFTFCGSFTGWRGHAHTYLTML

LSIAIWAVITLLMLPDDRRWDDTILSSALAANGVFLLAYVSPFWLLTKORNPMD

YPVEDAFCKPQLVKKSYGVENRAYSQEETIQGFEETGDTLYAPYSTHFQLOQPPOKE

FSIPRAHAWPSPKDYEVKKEGS"

QY 1 GCNCAYGCGTGGCCNTCNCNTAYARGAYTAYGARGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

Db 1271 GCCCAGCGCTGGCGAGCCCTTACAAGACTATGAAGTAAAG 1312

RESULT 14

AK001761

LOCUS 2446 bp mRNA linear PRI 20-JAN-2006

DEFINITION Homo sapiens cDNA FLJ10899 fis, clone NT2RP5003506.

ACCESSION AK001761

VERSION AK001761.1 GI:7023229

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominiidae; Homo.

1

Ota, T.,

Suzuki, Y.,

Nishikawa, T.,

Otsuki, T.,

Sugiyama, T.,

Irie, R.,

Wakamatsu, A.,

Hayashi, K.,

Sato, H.,

Nagai, K.,

Kimura, K.,

Makita, H.,

Sekine, M.,

Obayashi, M.,

Nishi, T.,

Shibahara, T.,

Tanaka, T.,

Ishii, S.,

Yamamoto, J.,

Saito, K.,

Kawai, Y.,

Isono, Y.,

Nakamura, Y.,

Nagabari, K.,

Murakami, K.,

Yasuda, T.,

Iwayanagi, T.,

Wagatsuma, M.,

Shiratori, A.,

Sudo, H.,

Hosoi, T.,

Kaku, Y.,

Kodaira, H.,

Kondo, H.,

Sugawara, M.,

Takahashi, M.,

Kanda, K.,

Tokoi, T.,

Furuya, T.,

Kikkawa, S.,

Omura, Y.,

Abé, K.,

Kamihara, K.,

Katsuta, N.,

Sato, K.,

Tanikawa, M.,

Ninomiya, K.,

Ishibashi, T.,

Yamashita, H.,

Murakawa, K.,

Fujimori, K.,

Tanai, H.,

Kimata, M.,

Watanabe, S.,

Hiraoka, S.,

Chiba, Y.,

Ishida, S.,

Ono, Y.,

Taguchi, S.,

Watanabe, S.,

Yosida, M.,

Hotuta, T.,

Kusano, J.,

Kanehori, K.,

Takahashi, T.,

Fuji, A.,

Hara, H.,

Tanase, T.,

Omura, Y.,

Togiya, S.,

Komai, F.,

Hara, R.,

Takeuchi, K.,

Arita, M.,

Imose, N.,

Musashino, K.,

Yuuki, H.,

Oshima, A.,

Sasaki, N.,

Aotsuka, S.,

Yoshikawa, Y.,

Matsunawa, H.,

Ichihara, T.,

Shiohata, N.,

Sano, S.,

Moriya, S.,

Moniyama, H.,

Sato, N.,

Takami, S.,

Goto, Y.,

Shimizu, F.,

Wakebe, H.,

Hishigaki, H.,

Watanabe, T.,

Sugiyama, A.,

Itakemoto, M.,

Kawakami, S.,

Yamamori, Y.,

Komiyama, K.,

Kumagai, A.,

Itakura, S.,

Fukuzumi, Y.,

Fujimori, Y.,

Ono, T.,

Yamada, K.,

Fujii, Y.,

Tashiro, H.,

Tanigami, A.,

Fujiwara, T.,

Ono, T.,

Yamada, K.,

Fujii, Y.,

Ozaki, K.,

Hirao, M.,

Ohmori, Y.,

Kawabata, A.,

Hikiji, T.,

Kobatake, N.,

Inagaki, H.,

Ikema, Y.,

Okamoto, S.,

Okitani, R.,

Kawakami, T.,

Noguchi, S.,

Itoh, T.,

Shigeta, K.,

Senba, T.,

Matsumura, K.,

Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

2
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Kanehori,K.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2446)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flj-cdna@nifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
Location/Qualifiers
1. .2446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP5003506"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP5"
/notes="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
254. .1327
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAA91890.1"
/db_xref="GI:7023230"
/translation="MATTVPDGRNGLSKYYRLCDKAEAWGIVLETVAAGVTVSA
FMLTPLIVCKVQDSNRKMLPTQFLFLGLVGLFGLTFAFIIGLDSTGPTFFLFG
ILFSCIFCLAHAVSLTKLVGRKPLSLVGLAVGSLVDVIAIEVILVMT
NNVFSLSAPRNEQFVLLTYVLFMLALTFLMSSFTFCGSGFTGKRGHGLVLTML
LSIAIWNATLMLDFDRRWDITLSSALAANGVFLLAYVSPFWLLTKORPMH
YPVEDAFCKPOLVKKSGVSNRAYSOEBEITQGFETGDTLYAPSTHFQLONPPOKE
FSIPRAHWSPYKDYEVKKEGS"

CDS
Query Match 90.4%; Score 30; DB 5; Length 2446;
Best Local Similarity 64.3%; Pred. No. 0.0085;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

ORIGIN
1 GCNCAYCGNTGGCCNTCNCNTAYAAAGATYAGACTGAAGTAAAG 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1271 GCCCAGCGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 15
LOCUS CQ981495 2456 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 350 from Patent EP1498424.
ACCESSION CQ981495
VERSION CQ981495.1 GI:58190785
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Rosenthal,A., Hermann,K., Heiden,E., Pilarsky,C., Bruemendorf,T.,
Staub,E., Roepcke,S., Mennerich,D., Kinnemann,H. and Li,X.
Human nucleic acid sequences from lung tumours
Patent: EP 1498424-A 350 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE);
Rosenthal, Andre (DE)
FEATURES
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0085;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCGNTGGCCNTCNCNTAYAAAGATYAGACTGAAGTAAAG 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1271 GCCCAGCGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 16
LOCUS DD210040 2456 bp DNA linear PAT 19-JAN-2006
DEFINITION Methods of Diagnosis of Cancer, Compositions and Methods of
Screening for Modulators of Cancer.
ACCESSION DD210040.1 GI:85654022
KEYWORDS JP 2005518782-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2456)
Zlotnik,A., Mack,D.H., Agiz,N., Gish,K.C., Hebeji,P.A., Wilson,K.E.
and Afar,D.
TITLE Methods of Diagnosis of Cancer, Compositions and Methods of
Screening for Modulators of Cancer
JOURNAL Patent: JP 2005518782-A 34 30-JUN-2005;
COMMENT PROTEIN DESIGN LABS INC
OS Homo sapiens
PN JP 2005518782-A/34
PD 30-JUN-2005
PF 17-SEP-2002 JP 2003529912
PR 12-APR-2002 US 60/372246,08-FEB-2002 US 60/355257, PR
08-FEB-2002 US 60/355145,13-NOV-2001 US 60/350666, PR
20-FEB-2001 US 60/323887,17-SEP-2001 US 60/323469 PI albert
zlotnik,david h mack,natasha agiz,kurt c gish,peter a pi hebeji,
pi keith e wilson,daniel afar
CC
FH Key Location/Qualifiers.
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source
1. .2456
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0085;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCGNTGGCCNTCNCNTAYAAAGATYAGACTGAAGTAAAG 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1271 GCCCAGCGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

JOURNAL Patent: WO 0142467-A 4043 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
source 1..3371
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 3371;
Best Local Similarity 64.3%; Pred. No. 0.009;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYGCNTGGCCNTCNCNTAYARGAYTAYGAGTNAAR 42
Db 1184 GCCCAGCGTTGGCGGACCCCTTACAAAGACTATGAGTAAG 1225
RESULT 20
CQ894732
LOCUS 6730 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 42 from Patent EP1471075.
ACCESSION CQ894732
VERSION CQ894732.1 GI:55467481
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Rosenthal, A.D., Pilarsky, C., Dahl, E., Specht, T., Bruemmendorf, T.,
Lichtner, R., Staub, E., Roepcke, S. and Li, X.I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: EP 1471075-A 42 27-OCT-2004;
Hinzmann, Bernd (DE); Rosenthal, Andre (DE); Pilarsky, Christian
(DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie
(DE)
FEATURES Location/Qualifiers
source 1..6730
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 6730;
Best Local Similarity 64.3%; Pred. No. 0.01;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYGCNTGGCCNTCNCNTAYARGAYTAYGAGTNAAR 42
Db 1271 GCCCAGCGTTGGCGGACCCCTTACAAAGACTATGAGTAAG 1312
RESULT 21
AC007688/c
LOCUS 161577 bp DNA linear PRI 30-AUG-2002
DEFINITION Homo sapiens 12 BAC RPC111-392P7 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
ACCESSION AC007688
VERSION AC007688.15 GI:5815499
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 161577)
AUTHORS Muzny, D., Arenson, A.D., Bouck, J., Bunac, C., Chen, Z., Ding, Y.,
Dugan, S., Durbin, J., Forcum, J., Garcia, C., Gorrell, J.H.,
Gorrell, L.L., Hernandez, J., Issar, A., Jackson, L., Kneitz, S.,
Kondejewski, N., Lau, S., Leal, B., Lee, E., Lichtarge, O., Liu, W.,
Logan, O., Lu, J., Marondel, I., Martinez, C., Merscher, S., Miller, A.,
Montgomery, K., Oswal, G., Pampell, L.R., Parish, B.J., Perez, L.,
Rashid, N.D., Rives, C., Scherer, S.E., Shen, H., Shim, C., Simon, M.,
Vo, Q., Williamson, A., Worley, K.C., Xhang, A.M., Yang, R., Yu, W.,
Zhou, X., Kucherlapati, R., Nelson, D. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 161577)
Worley, K.C.
Direct Submission
Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 1, 1999 this sequence version replaced gi:5757565.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.
ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.
SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.
QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.
QUALSTAT-REPORT-----
----- Summary Statistics -----
Contig length: 161577
Phrap values in estimate: 160751

Average error rate (BCM-Phrap estimate):0.000163681

Fraction of Phrap values less than 40 :0.0376047

Number of consensus changing edits:30

Number of N's in consensus :0

----- Consensus changing edits -----

Position

Original+Context

Edited+Context

7033

acctgcgcgt(n)ccgcgccctt

acctgcgcgt(c)ccgcgccctt

47567

aaaaaanaa(n)ggaataaat

aaaaaanaa(a)ggaataaat

51135

aaagaaga(n)aaagaagaa

aaagaaga(a)aaagaagaa

75582

aaaaaanaa(n)aangagatgt

aaaaaanaa(a)aaagatgtt

75585

aaaaaanaa(n)ggaagtctgc

aaaaaanaa(a)ggaagtctgc

75667

ctaaagcaga(n)taagatctta

ctaaagcaga(a)taagatctta

75752

ttttaaatag(n)gctttgcttt

ttttaaatag(t)gctttgcttt

84017

ggaggggaag(n)aaggaagag

ggaggggaag(g)aaggaagag

85227

tttgttttt(n)tgttttttt

tttgttttt(t)tgttttttt

92681

ggagttcaag(n)atgcagtgag

ggagttcaag(g)atgcagtgag

99681

aatctctat(n)ccgaattca

aatctctat(g)ccgaattca

111307

actagatatac(n)atccttttt

actagatatac(t)atccttttt

111374

gtatacnatc(n)ttttttttt

gtatacnatc(c)ttttttttt

111378

aaaaaanaa(n)ccatctctaga

aaaaaanaa(a)ccatctctaga

112621

atcacacctc(c)ttttttttt

atcacacctc(t)ttttttttt

135812

ttgcaggcac(n)cgccaccag

ttgcaggcac(a)cgccaccag

137207

cgccaccag(n)ctggctaagt

cgccaccag(c)ctggctaagt

137218

gcaagtga(n)natgtagaat

gcaagtga(a)natgtagaat

145113

caagtgaan(n)atgtagaata

caagtgaan(g)atgtagaata

145114

ggcacgggtg(n)ntcacnccag

ggcacgggtg(c)ntcacnccag

145232

ggcacgggtg(n)tcacnccag

ggcacgggtg(c)tcacnccag

145233

ggtgnntcac(n)ccagtaatcc

ggtgnntcac(g)ccagtaatcc

145321

gcaacatggt(t)nuncctatc

gcaacatggt(g)aaacccatc

145322

caacatggt(n)nncctatct

caacatggt(g)aaacccatct

145323

aacatggttn(n)ncctatctc

aacatggttn(a)acccatctc

145324

acatggttnn(n)cccatctct

acatggttnn(a)acccatctct

145367

gttcaccagg(n)gtggtggcgt

gttcaccagg(t)gtggtggcgt

145377

gtggcgta(n)tgtgagtcct

gtggcgta(g)tgtgagtcct

145397

tagctacttg(n)gagggagga

tagctacttg(g)gagggagga

145505

aaaaaaaa(n)gaaaaaaa

aaaaaaaa(a)gaaaaaaa

----- Distribution of Quality < 40 Bases -----

#

1000

*

900

*

800

*

700

*

bases

600

*

500

*

400

*

300

*

200

*

100

*

0

*

----- Phrap Value Range -----

5

*

10

*

15

*

20

*

25

*

30

*

35

*

40

*

Version: 1.01 qxfo.

Location/Qualifiers

1. .161577

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="12"

/clone="RP11-392P7"

complement(3...65)

/rpt_family="MER4B"

161..471

/rpt_family="AluXs"

complement(607..702)

/rpt_family="MIR"

FEATURES

source

repeat_region

repeat_region

repeat_region

repeat_region	703..838	/rpt_family="FLAM C"
repeat_region	complement(839..916)	/rpt_family="MIR"
repeat_region	complement(1205..1377)	/rpt_family="MER104"
STS	1388..1525	/standard_name="WIAF-759-STS"
	/db_xref="dbSTS:65358"	
repeat_region	1683..1971	/rpt_family="AluJo"
repeat_region	1974..2261	/rpt_family="AluJb"
repeat_region	2486..2593	/rpt_family="MIR"
misc_feature	2610..3214	/note="Region: Unigene cluster similar to AA056332 and A1074576"
STS	2712..2823	/standard_name="SHGC-44583"
	/db_xref="dbSTS:48451"	
repeat_region	3185..3491	/rpt_family="AluSq"
repeat_region	complement(6397..6537)	/rpt_family="L2"
repeat_region	6967..7158	/rpt_family="(CCCCG)n"
repeat_region	7094..7240	/rpt_family="(CG)n"
repeat_region	complement(7770..8250)	/rpt_family="MER44B"
repeat_region	8831..9073	/rpt_family="MER102"
repeat_region	complement(9075..9303)	/rpt_family="L2"
repeat_region	9671..10135	/rpt_family="LTR33A"
Query Match	90.4%;	Score 30; DB 5; Length 161577;
Best Local Similarity	64.3%;	Pred. No. 0.019;
Matches	27; Conservative	7; Mismatches 8; Indels 0; Gaps 0;
Qy	1	GCNCAAGTGGCCNTCNCNTAYARGAYTARGGTNAAR 42
Db	94911	GCCACGCTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 94870
RESULT 22		
BD150346		
LOCUS		
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.	
ACCESSION	BD150346	
VERSION	BD150346.1 GI:27856104	
KEYWORDS	JP 2002191363-A/5189.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 680)	
AUTHORS	Ota T., Isogai T., Nishikawa T., Hayashi K., Saito K., Yamamoto J., Ishii S., Sugiyama T., Wakamatsu A., Nagai K. and Otsuki T.	
TITLE	Primer for synthesizing full-length cDNA and use thereof	
JOURNAL	Patent: JP 2002191363-A 5189 09-JUL-2002;	
	HELIX RESEARCH INSTITUTE	
COMMENT	OS Homo sapiens (human)	
	EN JP 2002191363-A/5189	
	PD 09-JUL-2002	
	PF 28-JUL-2000 JP 2000280990	
	PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,	
	PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,	
	PI KEIICHI NAGAI, TETSUJI OTSUKI	


```

PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10, C12N21/02,C12N1/68/C12P21/08,C06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source
1. .680
/organism='Homo sapiens (human)'.
FEATURES
source
1..680
Location/Qualifiers
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 82.5%; Score 27.4; DB 2; Length 680;
Best Local Similarity 59.5%; Pred. No. 0.13;
Matches 25; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGATYAGTGTAAR 42
|||:||||| |||:|||||:||||: |||:
Db 633 GCCCAGCTTGGCCCAACCTTACAAAGACTATGAANGTAA 674

RESULT 23
AX870284 680 bp DNA linear PAT 17-DEC-2003
LOCUS
DEFINITION Sequence 5189 from Patent EP1074617.
ACCESSION AX870284
VERSION AX870284.1 GI:40025147
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 5189 07-FEB-2001;
Research Association for Biotechnology (JP)
Location/Qualifiers
FEATURES
source
1..680
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 82.5%; Score 27.4; DB 2; Length 680;
Best Local Similarity 59.5%; Pred. No. 0.13;
Matches 25; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGATYAGTGTAAR 42
|||:||||| |||:|||||:||||: |||:
Db 633 GCCCAGCTTGGCCCAACCTTACAAAGACTATGAANGTAA 674

RESULT 24
AC149246 224653 bp DNA linear HTG 19-JUN-2004
LOCUS
DEFINITION Orolemur garnettii clone CH256-3707, WORKING DRAFT SEQUENCE, 8
ordered pieces.
ACCESSION AC149246
VERSION AC149246.2 GI:48958698
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Orolemur garnettii (small-eared galago)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Strepsirhini;
Galagonidae; Orolemur.
REFERENCE
1 (bases 1 to 224653)
AUTHORS Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,

```

```

Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
Karlsn,E., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q.,
Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
Masiello,C., Maskeri,B., McDowell,J., Mullikin,J.C., Paquirigan,C.,
Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddik-Dugue,N.,
Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantiripop,S.,
Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D.,
Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 224653)
Green,E.D.
Direct Submission
Submitted (27-MAY-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 224653)
Green,E.D.
Direct Submission
Submitted (19-JUN-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Jun 19, 2004 this sequence version replaced gi:47716538.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hgri.nih.gov
----- Project Information
Center project name: gpg
Center clone name: 037007

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 223254 bases at least Q40
Consensus quality: 223828 bases at least Q30
Consensus quality: 223915 bases at least Q20
Insert size: 233000; agarose-fp
Insert size: 223953; sum-of-contigs
Quality coverage: 11.53x in Q20 bases; agarose-fp
Quality coverage: 12.00x in Q20 bases; sum-of-contigs
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

```

* 1 8440: contig of 8440 bp in length
* 8441 21103: contig of 12563 bp in length
* 8541 21203: gap of unknown length
* 21104 36328: contig of 15125 bp in length
* 21204 36429: gap of unknown length
* 36329 45563: contig of 9135 bp in length
* 36429 45564 45663: gap of unknown length

```



```

REFERENCE
AUTHORS      Sulston,J.E. and Waterston,R.
TITLE        Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
PUBMED      9847074

REFERENCE
AUTHORS      Wallgoriski,J., Haglund,K. and Creason,K.
TITLE        The sequence of Homo sapiens BAC clone RP11-732B24
JOURNAL      Unpublished (2001)

REFERENCE
AUTHORS      Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (10-SEP-2001) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA

REFERENCE
AUTHORS      Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (12-OCT-2001) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA

REFERENCE
AUTHORS      Waterston,R.
TITLE        Direct Submission
JOURNAL      Submitted (01-MAR-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Oct 12, 2001 this sequence version replaced gi:15778802.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc
              Contact: sapiens@watson.wustl.edu
              ----- Summary Statistics
              Center project name: H_NH0732B24
              Drafting Center: WIBR
              -----

NOTE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanesi,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-401119; the clone sequenced
to the right is RP11-48814. Actual start of this clone is at base
position 1 of RP11-732B24; actual end is at base position 179591 of
RP11-732B24.

```

Data from AC025554 was used to finish this clone, AC093902.
Polymorphisms have been identified between AC025554 and AC093902.

The sequence of AC026505 has been incorporated into AC093902.

```

FEATURES
source
1..179591
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-732B24"
/clone_lib="RPCI-11"
997..1710
/notes="similar to EST BI460037 (NID:gl5250693)"
repeat_region
2534..2622
/rpt_family="MER1_type"
2751..2865
/rpt_family="MER1_type"
3341..3676
/notes="similar to EST T84790 (NID:g713142) yd51f01.r1"
misc_feature
3349..3844
/notes="similar to EST T72623 (NID:g689298) yd19b05.s1"
repeat_region
3558..3859
/rpt_family="Alu"
repeat_region
4004..4018
/rpt_family="L1"
repeat_region
4019..4165
/rpt_family="L1"
repeat_region
4270..4383
/rpt_family="(TA)n"
repeat_region
4413..4502
/rpt_family="MER99"
repeat_region
4921..5061
/rpt_family="L1"
repeat_region
6148..6197
/rpt_family="AT_rich"
repeat_region
6393..6415
/rpt_family="AT_rich"
repeat_region
6581..6719
/rpt_family="AT_rich"
repeat_region
6791..6828
/rpt_family="MIR"
repeat_region
6869..7367
/rpt_family="AT_rich"
repeat_region
7416..7499
/rpt_family="L2"
repeat_region
7632..7821
/rpt_family="L1"
repeat_region
8235..8281
/rpt_family="(TA)n"
repeat_region
8507..8543
/rpt_family="AT_rich"
repeat_region
8511..8916
/rpt_family="L1"
repeat_region
8643..8670
/rpt_family="AT_rich"
repeat_region
8932..9007
/rpt_family="ERV1"
repeat_region
9008..9184
/rpt_family="L1"
repeat_region
9185..9578
/rpt_family="MALR"
repeat_region
9370..9393
/rpt_family="(GA)n"
repeat_region
9579..12312
/rpt_family="L1"
repeat_region
12307..12331
/rpt_family="AT_rich"
repeat_region
12337..12971
/rpt_family="L1"
repeat_region
12730..12766
/rpt_family="AT_rich"

```

```
repeat_region 12977..13263
/rpt_family="L2"
repeat_region 13510..13532
/rpt_family="AT_rich"
repeat_region 13603..13691
/rpt_family="L1"
repeat_region 13895..13920
/rpt_family="AT_rich"
repeat_region 14918..14951
/rpt_family="(T)n"
repeat_region 14922..15232
/rpt_family="Alu"
repeat_region 15697..15719
/rpt_family="AT_rich"
repeat_region 18777..18841
/rpt_family="MIR"
repeat_region 18881..18913
/rpt_family="(CATATA)n"
repeat_region 19141..19163
/rpt_family="AT_rich"
repeat_region 19255..19285
/rpt_family="AT_rich"
repeat_region 19381..19436
/rpt_family="(TA)n"
repeat_region 19771..19794
/rpt_family="AT_rich"
repeat_region 19974..20278
/rpt_family="Alu"
repeat_region 20252..20286
/rpt_family="(CAAAA)n"
repeat_region 20413..20478
/rpt_family="AT_rich"
repeat_region 20578..20697
/rpt_family="L1"
repeat_region 20698..20913
/rpt_family="Alu"
repeat_region 20892..20926
/rpt_family="(TAAA)n"
repeat_region 20914..21078
/rpt_family="L1"
repeat_region 21110..21314
/rpt_family="L1"
repeat_region 21121..21218
/rpt_family="A-rich"
repeat_region 21273..21295

Query Match 65.1%; Score 21.6; DB 5; Length 179591;
Best Local Similarity 56.8%; Pred. No. 2.6e+02;
Matches 21; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 6 YGNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
:|||||:|||||:|||||:|||||:|||||:
Db 108898 TGCTGGCTGCTCCCTACGAGGATGAAGTAAAG 108934

RESULT 29
AL606745/c
LOCUS Mouse DNA sequence from clone RP23-32L6 on chromosome 3, complete
DEFINITION sequence.
ACCESSION AL606745.11 GI:24527386
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 192332)
Unpublished
High Throughput Mouse Sequencing
2 (bases 1 to 203317)
Grills, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Toshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A., Gordon, M.
and Kucherlapati, R.
Direct Submission
Submitted (01-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

repeat_region 12977..13263
/rpt_family="L2"
repeat_region 13510..13532
/rpt_family="AT_rich"
repeat_region 13603..13691
/rpt_family="L1"
repeat_region 13895..13920
/rpt_family="AT_rich"
repeat_region 14918..14951
/rpt_family="(T)n"
repeat_region 14922..15232
/rpt_family="Alu"
repeat_region 15697..15719
/rpt_family="AT_rich"
repeat_region 18777..18841
/rpt_family="MIR"
repeat_region 18881..18913
/rpt_family="(CATATA)n"
repeat_region 19141..19163
/rpt_family="AT_rich"
repeat_region 19255..19285
/rpt_family="AT_rich"
repeat_region 19381..19436
/rpt_family="(TA)n"
repeat_region 19771..19794
/rpt_family="AT_rich"
repeat_region 19974..20278
/rpt_family="Alu"
repeat_region 20252..20286
/rpt_family="(CAAAA)n"
repeat_region 20413..20478
/rpt_family="AT_rich"
repeat_region 20578..20697
/rpt_family="L1"
repeat_region 20698..20913
/rpt_family="Alu"
repeat_region 20892..20926
/rpt_family="(TAAA)n"
repeat_region 20914..21078
/rpt_family="L1"
repeat_region 21110..21314
/rpt_family="L1"
repeat_region 21121..21218
/rpt_family="A-rich"
repeat_region 21273..21295

Query Match 65.1%; Score 21.6; DB 6; Length 192332;
Best Local Similarity 56.8%; Pred. No. 2.6e+02;
Matches 21; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 2 CMCAYGCTGGCCNTCCNTAYARGAYTAYGARGT 38
:|||||:|||||:|||||:|||||:|||||:
Db 172227 CACACATGGCGCTACCCACACAGGCTAATGAAGT 172191

RESULT 30
AC021062/c
LOCUS Mus musculus strain C57BL6/J chromosome 3 clone RP23-102D8, WORKING
DEFINITION DRAFT SEQUENCE, 22 unordered pieces.
ACCESSION AC021062
VERSION AC021062.16 GI:21358691
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 203317)
Grills, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Toshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A., Gordon, M.
and Kucherlapati, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 203317)
Grills, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Toshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A., Gordon, M.
and Kucherlapati, R.
Direct Submission

COMMENT
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 3, 2002 this sequence version replaced gi:24394907.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-32L6 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES
source
1..192332
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/clone="RP23-32L6"
/clone_lib="RPCI-23"

ORIGIN
Query Match 65.1%; Score 21.6; DB 6; Length 192332;
Best Local Similarity 56.8%; Pred. No. 2.6e+02;
Matches 21; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 2 CMCAYGCTGGCCNTCCNTAYARGAYTAYGARGT 38
:|||||:|||||:|||||:|||||:|||||:
Db 172227 CACACATGGCGCTACCCACACAGGCTAATGAAGT 172191

RESULT 30
AC021062/c
LOCUS Mus musculus strain C57BL6/J chromosome 3 clone RP23-102D8, WORKING
DEFINITION DRAFT SEQUENCE, 22 unordered pieces.
ACCESSION AC021062
VERSION AC021062.16 GI:21358691
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 203317)
Grills, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Toshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A., Gordon, M.
and Kucherlapati, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 203317)
Grills, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Toshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A., Gordon, M.
and Kucherlapati, R.
Direct Submission
```

JOURNAL

Submitted (14-JAN-2000) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave., Bronx, NY 10461, USA
3 (bases 1 to 203317)

REFERENCE
AUTHORS

Grills, G., Han, J., J., Long, J., Pomerantz, R.,
Ioehikhes, I. P., Shim, C., Decker, J., Thomas, E., Perera, A., Gordon, M.
and Kucherlapati, R.

TITLE
JOURNAL

Direct Submission
Submitted (08-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA

COMMENT

On Jun 8, 2002 this sequence version replaced gi:21306613.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpcgg.org/Sequence/mouse.html>

Contact: hpgc@medel.mgh.harvard.edu

-----Summary Statistics

Center project name: ANY

Sequencing vector: pUC18; L08752

Chemistry: Dye-terminator Big Dye; 100%

*Consensus quality: 194428 at least Q20

*Consensus quality: 191055 at least Q30

*Consensus quality: 185439 at least Q40

Estimated insert size: agarose-FP - N/A

Quality coverage: 202897 - sum-of-contigs

Quality coverage: agarose-FP - N/A

Quality coverage: 4.3 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 24854: contig of 24854 bp in length
* 24855 24874: gap of unknown length
* 24875 45360: contig of 20486 bp in length
* 45361 45380: gap of unknown length
* 45381 62264: contig of 16884 bp in length
* 62265 62284: gap of unknown length
* 62285 81433: contig of 19149 bp in length
* 81434 81453: gap of unknown length
* 81454 97133: contig of 15680 bp in length
* 97134 97153: gap of unknown length
* 97154 111163: contig of 14010 bp in length
* 111164 111183: gap of unknown length
* 11184 125606: contig of 14422 bp in length
* 125606 125625: gap of unknown length
* 125626 137365: contig of 11739 bp in length
* 137365 137384: gap of unknown length
* 137385 149894: contig of 12510 bp in length
* 149895 149915: gap of unknown length
* 149915 158741: contig of 8826 bp in length
* 158741 158761: gap of unknown length
* 158761 166785: contig of 8025 bp in length
* 166786 166806: gap of unknown length
* 166806 172697: contig of 5892 bp in length
* 172698 172717: gap of unknown length
* 172718 176174: contig of 3457 bp in length
* 176175 176194: gap of unknown length
* 176195 181319: contig of 5124 bp in length
* 181319 181338: gap of unknown length
* 181339 185745: contig of 4407 bp in length
* 185746 185766: gap of unknown length
* 185766 191167: contig of 5401 bp in length
* 191167 191187: gap of unknown length
* 191187 195048: contig of 3861 bp in length
* 195048 195068: gap of unknown length
* 195068 197304: contig of 2237 bp in length
* 197305 197324: gap of unknown length

* 197325 198964: contig of 1640 bp in length
* 198965 198984: gap of unknown length
* 198985 200466: contig of 1482 bp in length
* 200467 200486: gap of unknown length
* 200487 201931: contig of 1445 bp in length
* 201932 201951: gap of unknown length
* 201952 203317: contig of 1366 bp in length.
FEATURES
source
1. .203317
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="3"
/clone="RP23-102D8"
/sex="male"
1. .24854
/note="assembly_name:Contig51"
24855..24874
/estimated_length=unknown
24875..45360
/note="assembly_name:Contig50"
45361..45380
/estimated_length=unknown
45381..62264
/note="assembly_name:Contig49"
62265..62284
/estimated_length=unknown
62285..81433
/note="assembly_name:Contig48"
81434..81453
/estimated_length=unknown
81454..97133
/note="assembly_name:Contig47"
97134..97153
/estimated_length=unknown
97154..111163
/note="assembly_name:Contig46"
111164..111183
/estimated_length=unknown
111184..125605
/note="assembly_name:Contig45"
125606..125625
/estimated_length=unknown
125626..137364
/note="assembly_name:Contig44"
137365..137384
/estimated_length=unknown
137385..149894
/note="assembly_name:Contig43
clone_end:T7
vector_side:right"
149895..149914
/estimated_length=unknown
149915..158740
/note="assembly_name:Contig42"
158741..158760
/estimated_length=unknown
158761..166785
/note="assembly_name:Contig41
clone_end:SP6
vector_side:left"
166786..166805
/estimated_length=unknown
166806..172697
/note="assembly_name:Contig40"
172698..172717
/estimated_length=unknown
172718..176174
/note="assembly_name:Contig39"
176175..176194
/estimated_length=unknown
176195..181318
/estimated_length=unknown

```

/note="assembly_name:Contig38"
181319..181338
/estimated_length=unknown
181339..185745
/note="assembly_name:Contig37"
185746..185765
/estimated_length=unknown
185766..191166
/note="assembly_name:Contig36"
191167..191186
/estimated_length=unknown
191187..195047
/note="assembly_name:Contig35"
195048..195067
/estimated_length=unknown
195068..197304
/note="assembly_name:Contig34"
197305..197324
/estimated_length=unknown
197325..198964
/note="assembly_name:Contig33"
198965..198984
/estimated_length=unknown
198985..200466
/note="assembly_name:Contig32"
200467..200486
/estimated_length=unknown
200487..201931
/note="assembly_name:Contig31"
201932..201951
/estimated_length=unknown
201952..203317

Query Match      65.1%; Score 21.6; DB 12; Length 203317;
Best Local Similarity 56.8%; Pred. No. 2.7e+02;
Matches 21; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY  2  CNYAGGNTGGCCNTCCNTTAYARGATYTAGGT 38
Db   78906  CACACATGGCCCTACCCACACAGGCTAATGAGT 78970

RESULT 31
AK111295
LOCUS      AK111295      859 bp      mRNA      linear      PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:002-181-A09, full
            insert sequence.
ACCESSION  AK111295
VERSION    AK111295.1 GI:32996504
KEYWORDS   FLI CDNA; oligo capping.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  1
AUTHORS    The Rice Full-length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-length cDNA Project Team,
            Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
            Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
            Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
            Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group, Ohtsuki, K., Murakami, K.,
            Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
            Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, N., Xie, Q., Lu, M.,
            Nariikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
            Ikeda, R., Iehibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
            Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
            Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
            Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
            Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y.,
            Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
            Yoshino, M. and Hayashizaki, Y.

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
12869764
2 (bases 1 to 859)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Ohtsuki, K., Saitoh, H., Saitoh, K., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp).
Tel:81-29-838-7007. Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ohtsuki, K., Iida, Y.,
Fujimura, T., Ikeda, R., Iehibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Imotani, K.,
Ishii, Y., Itoh, M., Kawai, J., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Sugami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
FEATURES
source
1..859
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="002-181-A09"
ORIGIN
Query Match      64.5%; Score 21.4; DB 4; Length 859;
Best Local Similarity 53.7%; Pred. No. 1.2e+02;
Matches 22; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

```


AP008217_049	4900001	5010000	AP008217_122	12200001	12310000
AP008217_050	5000001	5110000	AP008217_123	12300001	12410000
AP008217_051	5100001	5210000	AP008217_124	12400001	12510000
AP008217_052	5200001	5310000	AP008217_125	12500001	12610000
AP008217_053	5300001	5410000	AP008217_126	12600001	12710000
AP008217_054	5400001	5510000	AP008217_127	12700001	12810000
AP008217_055	5500001	5610000	AP008217_128	12800001	12910000
AP008217_056	5600001	5710000	AP008217_129	12900001	13010000
AP008217_057	5700001	5810000	AP008217_130	13000001	13110000
AP008217_058	5800001	5910000	AP008217_131	13100001	13210000
AP008217_059	5900001	6010000	AP008217_132	13200001	13310000
AP008217_060	6000001	6110000	AP008217_133	13300001	13410000
AP008217_061	6100001	6210000	AP008217_134	13400001	13510000
AP008217_062	6200001	6310000	AP008217_135	13500001	13610000
AP008217_063	6300001	6410000	AP008217_136	13600001	13710000
AP008217_064	6400001	6510000	AP008217_137	13700001	13810000
AP008217_065	6500001	6610000	AP008217_138	13800001	13910000
AP008217_066	6600001	6710000	AP008217_139	13900001	14010000
AP008217_067	6700001	6810000	AP008217_140	14000001	14110000
AP008217_068	6800001	6910000	AP008217_141	14100001	14210000
AP008217_069	6900001	7010000	AP008217_142	14200001	14310000
AP008217_070	7000001	7110000	AP008217_143	14300001	14410000
AP008217_071	7100001	7210000	AP008217_144	14400001	14510000
AP008217_072	7200001	7310000	AP008217_145	14500001	14610000
AP008217_073	7300001	7410000	AP008217_146	14600001	14710000
AP008217_074	7400001	7510000	AP008217_147	14700001	14810000
AP008217_075	7500001	7610000	AP008217_148	14800001	14910000
AP008217_076	7600001	7710000	AP008217_149	14900001	15010000
AP008217_077	7700001	7810000	AP008217_150	15000001	15110000
AP008217_078	7800001	7910000	AP008217_151	15100001	15210000
AP008217_079	7900001	8010000	AP008217_152	15200001	15310000
AP008217_080	8000001	8110000	AP008217_153	15300001	15410000
AP008217_081	8100001	8210000	AP008217_154	15400001	15510000
AP008217_082	8200001	8310000	AP008217_155	15500001	15610000
AP008217_083	8300001	8410000	AP008217_156	15600001	15710000
AP008217_084	8400001	8510000	AP008217_157	15700001	15810000
AP008217_085	8500001	8610000	AP008217_158	15800001	15910000
AP008217_086	8600001	8710000	AP008217_159	15900001	16010000
AP008217_087	8700001	8810000	AP008217_160	16000001	16110000
AP008217_088	8800001	8910000	AP008217_161	16100001	16210000
AP008217_089	8900001	9010000	AP008217_162	16200001	16310000
AP008217_090	9000001	9110000	AP008217_163	16300001	16410000
AP008217_091	9100001	9210000	AP008217_164	16400001	16510000
AP008217_092	9200001	9310000	AP008217_165	16500001	16610000
AP008217_093	9300001	9410000	AP008217_166	16600001	16710000
AP008217_094	9400001	9510000	AP008217_167	16700001	16810000
AP008217_095	9500001	9610000	AP008217_168	16800001	16910000
AP008217_096	9600001	9710000	AP008217_169	16900001	17010000
AP008217_097	9700001	9810000	AP008217_170	17000001	17110000
AP008217_098	9800001	9910000	AP008217_171	17100001	17210000
AP008217_099	9900001	10010000	AP008217_172	17200001	17310000
AP008217_100	10000001	10110000	AP008217_173	17300001	17410000
AP008217_101	10100001	10210000	AP008217_174	17400001	17510000
AP008217_102	10200001	10310000	AP008217_175	17500001	17610000
AP008217_103	10300001	10410000	AP008217_176	17600001	17710000
AP008217_104	10400001	10510000	AP008217_177	17700001	17810000
AP008217_105	10500001	10610000	AP008217_178	17800001	17910000
AP008217_106	10600001	10710000	AP008217_179	17900001	18010000
AP008217_107	10700001	10810000	AP008217_180	18000001	18110000
AP008217_108	10800001	10910000	AP008217_181	18100001	18210000
AP008217_109	10900001	11010000	AP008217_182	18200001	18310000
AP008217_110	11000001	11110000	AP008217_183	18300001	18410000
AP008217_111	11100001	11210000	AP008217_184	18400001	18510000
AP008217_112	11200001	11310000	AP008217_185	18500001	18610000
AP008217_113	11300001	11410000	AP008217_186	18600001	18710000
AP008217_114	11400001	11510000	AP008217_187	18700001	18810000
AP008217_115	11500001	11610000	AP008217_188	18800001	18910000
AP008217_116	11600001	11710000	AP008217_189	18900001	19010000
AP008217_117	11700001	11810000	AP008217_190	19000001	19110000
AP008217_118	11800001	11910000	AP008217_191	19100001	19210000
AP008217_119	11900001	12010000	AP008217_192	19200001	19310000
AP008217_120	12000001	12110000	AP008217_193	19300001	19410000
AP008217_121	12100001	12210000	AP008217_194	19400001	19510000

```

AP008217_195      19500001 19610000
AP008217_196      19600001 19710000

Query Match      64.5%; Score 21.4; DB 4; Length 110000;
Best Local Similarity 53.7%; Pred. No. 3e+02;
Matches 22; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 CNCAYCGTCGCGCCGTCGCCATCTAAATATTCAGAGGTTTAG 53595
Db 53555 CGTCGCTCGCGCGTCGCCATCTAAATATTCAGAGGTTTAG 55244

RESULT 34
AC134924
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 11 BAC clone
OSJNB0076M06, complete sequence.
ACCESSION
AC134924
VERSION
AC134924.2 GI:38044142
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BCP
Clade; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 133581)
Gupta,V., Bharti,A.K., Raghuvanshi,S., *McCombie,W.R., *de la
Bastide,M., *Spiegel,L., *Zutavern,T., *Muller,S., *Nascimento,L.,
*Balijs,V., *Bell,M., *Miller,B., *Katzenberger,F., *Andrade,M.V.,
*Dike,S., *O'Shaughnessy,A., *Palmer,L., *Dedhi,N., Khurana,P.,
Khurana,J.P. and Tyagi,A.K.
Oryza sativa (japonica cultivar-group) chromosome 11 BAC clone
OSJNB0076M06, complete sequence
Unpublished
*Litte Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
2 (bases 1 to 133581)
Gupta,V., Bharti,A.K., Raghuvanshi,S., Khurana,P., Khurana,J.P. and
Tyagi,A.K.
Direct Submission
Submitted (02-OCT-2002) Indian Initiative for Rice Genome
Sequencing, Department of Plant Molecular Biology, University of
Delhi South Campus, New Delhi, Delhi 110021, India
3 (bases 1 to 133581)
Gupta,V., Bharti,A.K., Raghuvanshi,S., Khurana,P., Khurana,J.P. and
Tyagi,A.K.
Direct Submission
Submitted (30-OCT-2003) Indian Initiative for Rice Genome
Sequencing, Department of Plant Molecular Biology, Benito Juarez
Road, New Delhi, Delhi 110021, India
On Oct 30, 2003 this sequence version replaced gi:23463031.
A gap of 40 bp was filled by merging the data of contig
CL038074.171 from Syngenta draft sequence (Syd) of the rice genome.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The nucleotide
sequence of this BAC clone was sequenced to phase II by the Indian
Initiative for Rice Genome Sequencing and completed to phase III at
the Cold Spring Harbor Laboratory Genome Center under the Indo-USA
collaboration.
FEATURES
source
1. 133581
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica cultivar-group"
/db_xref="taxon:39947"
/chromosome="11"
/clone="OSJNB0076M06"

/clone_lib="EcoRI"
complement(1150..1382)
/note="The assembly is covered solely by reads which were
generated from a transposed plasmid subclone. The
sequence is high quality."
10075..10076
/note="We believe the assembly to be correct. However,
there are 2 subclone sequences which call only 1 G at this
position."
79547..79734
/note="The sequence is a simple sequence repeat (TAGA)n in
which the exact number of repeat units is unknown. Four
subclone sequences in the region show up to 7 additional
repeat units compared that which is represented by the
assembly."

misc_feature
Query Match      64.5%; Score 21.4; DB 4; Length 133581;
Best Local Similarity 53.7%; Pred. No. 3.1e+02;
Matches 22; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 CNCAYCGTCGCGCCGTCGCCATCTAAATATTCAGAGGTTTAG 55244
Db 55204 CGTCGCTCGCGCGTCGCCATCTAAATATTCAGAGGTTTAG 55244

RESULT 35
AL135927
LOCUS
DEFINITION
Human DNA sequence from clone RP11-54H19 on chromosome 1 contains
the 3' end of the LMNA gene for lamin A/C, the gene for a novel
protein similar to semaphorins (FLJ12287), a novel gene (KIAA0446),
the PMF1 gene for polyamine-modulated factor 1, the BGLAP gene for
bone gamma-carboxyglutamate (gla) protein (osteocalcin), the gene
for prostestin and adipon receptor family member VI (PAQR6) and the
3' end of the gene for Estlp-like protein B (EST1B), complete
sequence.
AL135927
AL135927.14 GI:10443354
KEYWORDS
HTG; BGLAP; EST1B; FLJ12287; KIAA0446; LMNA; PAQR6; Semaphorins.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 164168)
Howden,P.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Oct 1, 2000 this sequence version replaced gi:10185404.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep
This sequence was generated from part of bacterial clone contigs of
human chromosome 1, constructed by the Sanger Centre Chromosome 1
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,

```

except on the rare occasion of the clone being a YAC.
RP11-54H19 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6.

FEATURES

source	Location/Qualifiers	gene
	1. .164168 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="1" /clone="RP11-54H19" /clone_lib="RPCI-11.1"	gene
misc_feature	1	
gene	/notes="Clone left end: RP11-54H19" join(AL355388.30:176665..176864,2358..2813) /gene="LMNA"	mRNA
mRNA	/locus tag="RP11-54H19.1-012" join(AL355388.30:176665..176864,2358..2813) /gene="LMNA"	gene
gene	/locus tag="RP11-54H19.1-012" /product="lamin A/C" /notes="match: ESTs: Em:AI929169.1" join(2395..2962,18305..18461,22091..22216,22493..22663, 22875..23000,23589..23809,23902..24124,24609..24716, 24801..24920,25342..25431,26176..26445,26768..27777) /gene="LMNA"	gene
mRNA	/locus tag="RP11-54H19.1-001" join(2395..2962,18305..18461,22091..22216,22493..22663, 22875..23000,23589..23809,23902..24124,24609..24716, 24801..24920,25342..25431,26176..26445,26768..27777) /gene="LMNA"	gene
gene	/locus tag="RP11-54H19.1-001" /product="lamin A/C" /notes="match: ESTs: Em:AI872233.1 Em:BF035195.1 Em:BG420208.1 Em:BG701671.1 Em:BM468449.1 Em:BO940951.1 match: cDNAs: Em:AK004619.2 Em:AK026584.1 Em:BC014507.1 Em:BC018863.2 Em:M13452.1 Em:X03444.1" join(AL355388.30:176258..176343, AL355388.30:176628..176864,AL355388.30:184151..184262, 2395..2962,18305..18461,22091..22216,22493..22663, 22875..23000,23589..23809,23902..24124,24609..24716, 24801..24920,25342..25553) /gene="LMNA"	mRNA
mRNA	/locus tag="RP11-54H19.1-002" join(AL355388.30:176258..176343, AL355388.30:176628..176864,AL355388.30:184151..184262, 2395..2962,18305..18461,22091..22216,22493..22663, 22875..23000,23589..23809,23902..24124,24609..24716, 24801..24920,25342..25553) /gene="LMNA"	gene
gene	/locus tag="RP11-54H19.1-002" /product="lamin A/C" /notes="match: cDNAs: Em:AK056143.1 Em:X03445.1" join(2399..2577,2651..2962,18305..18461,22091..22216, 22493..22663,22875..23000,23589..23809,23902..24124, 24609..24716,24801..24920,25342..25431,26176..26445, 26768..26991) /gene="LMNA"	gene
mRNA	/locus tag="RP11-54H19.1-008" join(2399..2577,2651..2962,18305..18461,22091..22216, 22493..22663,22875..23000,23589..23809,23902..24124, 24609..24716,24801..24920,25342..25431,26176..26445, 26768..26991) /gene="LMNA"	gene
gene	/locus tag="RP11-54H19.1-008" /product="lamin A/C" /notes="match: cDNAs: Em:AK122732.1" join(AL355388.30:176602..176864,2401..2946) /gene="LMNA"	gene
mRNA	/locus tag="RP11-54H19.1-009" join(AL355388.30:176602..176864,2401..2946) /gene="LMNA"	gene

/locus tag="RP11-54H19.1-009"
/product="lamin A/C"
/notes="match: ESTs: Em:BQ431561.1"
join(AL355388.30:176295..176343,
AL355388.30:176628..176864,2401..2819)
/gene="LMNA"
/locus tag="RP11-54H19.1-011"
join(AL355388.30:176295..176343,
AL355388.30:176628..176864,2401..2819)
/gene="LMNA"
/locus tag="RP11-54H19.1-011"
/product="lamin A/C"
/notes="match: ESTs: Em:AUI30227.1"
join(2410..2962,18305..18461,22091..22216,22493..22663,
22875..23000,23589..23809,23902..24124,24609..24716,
24801..24920,25342..25431,26176..26295,26768..27005)
/gene="LMNA"
/locus tag="RP11-54H19.1-004"
join(2410..2962,18305..18461,22091..22216,22493..22663,
22875..23000,23589..23809,23902..24124,24609..24716,
24801..24920,25342..25431,26176..26295,26768..27005)
/gene="LMNA"
/locus tag="RP11-54H19.1-004"
/product="lamin A/C"
/notes="match: cDNAs: Em:AY357727.1"
join(2414..2962,18305..18461,22091..22216,22493..22663,
22875..23000,23589..23809,23902..24080,25391..25554)
/gene="LMNA"
/locus tag="RP11-54H19.1-003"
/product="lamin A/C"
/notes="match: cDNAs: Em:BC033088.1"
join(2607..2962,18305..18461,22091..22216,22493..22663,
22875..23000,23589..23809,23902..24124,24609..24716,
24801..24920,25342..25431,26176..26445,26768..26794)
/gene="LMNA"
/locus tag="RP11-54H19.1-008"
/standard_name="OTTHUMP00000015848"
/notes="match: proteins: Sw:P25545 Sw:P48678 Sw:P48679
Tr:Q91WF2 Tr:Q9DC21"
/codon_start=1
/product="lamin A/C"
/protein_id="CAI15522.1"
/db_xref="GI:55957498"
/db_xref="GOA:Q5TCJ2"
/db_xref="InterPro:IPR001322"
/db_xref="UniProtKB/TrEMBL:Q5TCJ2"
/translation="METPSQRATRSQAQASSTPLSPTRITRLQEKEDLOELNDRLAV
YIDRVSRLETAAGLRITSEEVSVSGIKAAEALGDKATLDSVAKERARL
QELSKVRFEFKELKARNTKEGDLIAAQRKLDALEALNSKEAALSTALSEKRLTG
ELHDLGRQVAKLAALGAQKQIQDEMLRVDAAENRLQTKKELDFOKNIYSEELRET
KRRLTGLVINDKQREFESRLADALQELRAEDQVOYKKELEKTSYAKLDNARQ
SAERNNLVGAHEELQOQSRIRIDSLSAQLSQKOLAAKEALRDLEDSIARERDTS
RRLAKEREMAEARMQOQLDEYQELLIDKLADNEIHYRKLLEGEERLRSPS
PYSQSRGRASSHSSTQGGSVTKRKLESTESRSFSQHARTSGVAVAEVDEEGK
FTDLNRKNQDSQGMNQIKRQNGDPLLTYPFPKFTLKAGQVTTIWAAGATHSP
PVLRLNKAQNTQCGNSLRATLNGTSEEVAMRKLVRSTVVEDEDEDGDLHLHHH
GHSCTSSGDPAEYLNLSRTVLCGTCCQPADKASGSGAQQVGGPISSGSSASSVTVTR
SVRSVGGSGSGFGDNLVTRSYLLGSSPRTQSPONCSIN"
join(2607..2962,18305..18461,22091..22216,22493..22663,
22875..23000,23589..23809,23902..24124,24609..24716,
24801..24920,25342..25431,26176..26295,26768..26794)
/gene="LMNA"
/locus tag="RP11-54H19.1-004"
/standard_name="OTTHUMP00000015845"
/notes="match: proteins: Tr:AA29466"
/codon_start=1
/product="lamin A/C"
/protein_id="CAI15523.1"


```
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 186979)
Worley,K.C.
Direct Submission
Submitted (04-MAR-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 186979)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 29, 2005 this sequence version replaced gi:60498786.
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FXRW
Center clone name: CH240-271C13
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 184378 bases at least Q40
Consensus quality: 185440 bases at least Q30
Consensus quality: 186154 bases at least Q20
Estimated insert size: 187412; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 76761: contig of 76761 bp in length
* 76762 76811: gap of 50 bp
* 76812 157228: contig of 80417 bp in length
* 157229 157322: gap of 94 bp
* 157323 193077: contig of 25755 bp in length
* 183078 183127: gap of 50 bp
* 183128 184628: contig of 1501 bp in length
* 184629 184728: gap of unknown length
* 184729 185871: contig of 1143 bp in length
* 185872 185971: gap of unknown length
* 185972 186979: contig of 1008 bp in length.
* Location/Qualifiers
* 1. .186979
* /organism="Bos taurus"
* /mol_type="genomic DNA"
* /db_xref="taxon:9913"
* /clone="CH240-271C13"
* 76762. .76811
*
FEATURES
source
gap
/estimated_length=50
157229. .157322
/estimated_length=94
183078. .183127
/estimated_length=50
184629. .184728
/estimated_length=unknown
185872. .185971
/estimated_length=unknown
ORIGIN
Query Match 64.5%; Score 21.4; DB 12; Length 186979;
Best Local Similarity 56.4%; Pred. No. 3.3e+02; Indels 0; Gaps 0;
Matches 22; Conservative 5; Mismatches 12;
2 CMCVCGNTGGCCNTCCNTAYARGAYTAYGARGTNA 40
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36519 CCCTTGATGGACTTCACCTCTAGAGATTTGATGCA 36557
|||
RESULT 38
AC152284
LOCUS
DEFINITION
AC152284
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Bos taurus
Bos taurus clone CH240-3J9, WORKING DRAFT SEQUENCE, 26 unordered
pieces.
AC152284
AC152284.3 GI:68227188
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Bos taurus (cattle)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 186244)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,W.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davalia,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowitz,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulsegad,H., Lozado,R.J., Lu,X., Ma,J.,
Maheehwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J.J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Flopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Fuazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
```


Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaana, D., Waldron, B., Walker, B., Wang, D.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 196244)
Worley, K.C.
Direct Submission
Submitted (01-NOV-2004) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 196244)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 26, 2005 this sequence version replaced gi:58038084.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FADY
Center clone name: CH240-3J9
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 187776 bases at least Q40
Consensus quality: 189933 bases at least Q30
Consensus quality: 191909 bases at least Q20
Estimated insert size: 190708; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
2579: contig of 2579 bp in length
2580: gap of 50 bp
18531: contig of 15902 bp in length
2630
18532: gap of 50 bp
18582: contig of 31055 bp in length
49637: gap of 50 bp
49687: gap of 50 bp
55880: contig of 6194 bp in length
55881: gap of 50 bp
55931: contig of 5513 bp in length
61444: gap of 50 bp

61494 65499: contig of 4006 bp in length
65549: gap of 50 bp
65550 67082: contig of 1533 bp in length
67083 67182: gap of unknown length
74137: contig of 6955 bp in length
74138 74187: gap of 50 bp
87010: contig of 12823 bp in length
87011 87060: gap of 50 bp
87061 101691: contig of 14631 bp in length
101741: gap of 50 bp
101742 104378: contig of 2637 bp in length
104379 104428: gap of 50 bp
104429 108843: contig of 4415 bp in length
108844 108893: gap of 50 bp
108894 125174: contig of 16281 bp in length
125175 125224: gap of 50 bp
125225 134633: contig of 9409 bp in length
134634 134683: gap of 50 bp
134684 143350: contig of 8667 bp in length
143351 143400: gap of 50 bp
143401 152324: contig of 8924 bp in length
152325 152374: gap of 50 bp
152375 166855: contig of 14481 bp in length
166856 166956: gap of unknown length
166957 167962: contig of 1007 bp in length
167963 168062: gap of unknown length
168063 169176: contig of 1114 bp in length
169177 169276: gap of unknown length
169277 170700: contig of 1424 bp in length
170701 172163: contig of 1363 bp in length
172164 172264: gap of unknown length
172265 174723: contig of 2460 bp in length
174724 174824: gap of unknown length
174825 176941: gap of unknown length
176942 178369: contig of 1428 bp in length
178370 178469: gap of unknown length
178470 182811: contig of 4342 bp in length
182812 182912: gap of unknown length
182912 196244: contig of 13333 bp in length.

FEATURES
source
1..196244
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-3J9"
2580..2629
/estimated_length=50
18532..18581
/estimated_length=50
49637..49686
/estimated_length=50
55881..55930
/estimated_length=50
61444..61493
/estimated_length=50
65500..65549
/estimated_length=50
67083..67182
/estimated_length=unknown
74138..74187
/estimated_length=50
87011..87060
/estimated_length=50
101692..101741
/estimated_length=50
104379..104428
/estimated_length=50
108844..108893
/estimated_length=50
125175..125224
/estimated_length=50

gap 134634..134683
/estimated_length=50
gap 143351..143400

Query Match 64.5%; Score 21.4; DB 12; Length 196244;
Best Local Similarity 53.7%; Pred.No.3.3e+02;
Matches 22; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 CMCAYGCGTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42
|||:| ||| ||| |||:| ||| |||:| ||| |||:
Db 155585 CTCATCCTCTGCCACCCCATACATATATTAGAAGTAAAA 155625

RESULT 39
AF110497 1604 bp DNA linear BCT 12-JAN-1999
LOCUS Bartonella quintana filament-A precursor (fila) gene, partial cds;
DEFINITION carboxy-terminal protease (ctpa) gene, complete cds; and
invasion-associated locus A protein (iala) gene, partial cds.
ACCESSION AF110497
VERSION AF110497.1 GI:4140713
KEYWORDS
SOURCE Bartonella quintana
ORGANISM Bartonella quintana
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bartonellaceae; Bartonella.
REFERENCE 1 (bases 1 to 1604)
AUTHORS Minnick,M.F., Smitherman,L.S. and Carroll,J.
TITLE A tail-specific protease from Bartonella quintana
JOURNAL Unpublished
2 (bases 1 to 1604)
AUTHORS Minnick,M.F., Smitherman,L.S. and Carroll,J.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1998) Division of Biological Sciences, The
University of Montana, Health Sciences 104, Missoula, MT
59812-1002, USA

FEATURES
source
1..1604
/organism="Bartonella quintana"
/mol_type="genomic DNA"
/strain="Houston 90-268"
/db_xref="taxon:803"
<1..203
/genes="fila"
<1..203
/genes="fila"
/notes="putative alpha-helical surface protein; similar to
Bartonella bacilliformis filament-A precursor"
/codon_start=3
/transl_table=11
/product="filament-A precursor"
/protein_id="AAD04179.1"
/db_xref="GI:4140715"
/translations="NVTQGFVLSGFLGTMGQFTANSVALDIGKTAPMLYIEFRKQ
GKPVNTPWQSEKTRQNDS"
193..1518
/genes="ctpa"
193..1518
/genes="ctpa"
/notes="tail-specific protease; similar to Bartonella
bacilliformis carboxy-terminal protease"
/codon_start=1
/transl_table=11
/product="carboxy-terminal protease"
/protein_id="AAD04178.1"
/db_xref="GI:4140714"
/translations="MRKVVILLVAGVLLGASSMIVQSVAAANNEGDTYQOLSFLFGDIF
ERVMDYIVPDDKKLIENAINGLTSLDPSHYMAEAKMDSTKGFEGGLGLEV
THEKNLVKVSPPMDPTPSKAGLAGLSDIGVGTNGQTLINEAVKMRGAGPTPI
LTIIRSGVDKPLEIKVVRDIIKVKAVRVVEDDIGYLRIRFSEQTFGDLQTAIKDIO
SKTPOKLVGLDLNPGLLNQAVNSAFLNKGEIVSTRKKNVDVVRDAPKG
DIINGPLVLINGSSASEIVAGALQDHRRTATILGTOSFGKGSVQTVIPLGECNAL
RUTALYYTPAGTSIQGTITPDIIIVEQPLPEQIKDYDVKLGESELKGHILKGQESNK

gene
CDS

GSGSAAFVPKDPKDDTQOLSEAYKLLRGEMAHAAFPDPNKNVLKQDVLK"
1560..>1604
/gene="iala"
1560..>1604
/gene="iala"
/note="similar to Bartonella bacilliformis
invasion-associated protein"
/codon_start=1
/transl_table=11
/product="invasion-associated locus A protein"
/protein_id="AAD04180.1"
/db_xref="GI:4140716"
/translation="WPTDINLXIFLIEDV"

ORIGIN

Query Match 63.9%; Score 21.2; DB 15; Length 1604;
Best Local Similarity 54.8%; Pred.No.1.7e+02;
Matches 23; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42
|||:| ||| ||| |||:| ||| |||:| ||| |||:
Db 1273 GAGCAACCATTCGCTGACATATATAAGATTACGATGTAAG 1314

RESULT 40
BX897700_01
WPCOMMENT
Sequence split into 16 fragments LOCUS BX897700 Accession BX897700

Fragment Name	Begin	End
BX897700_00	1	110000
BX897700_01	100001	210000
BX897700_02	200001	310000
BX897700_03	300001	410000
BX897700_04	400001	510000
BX897700_05	500001	610000
BX897700_06	600001	710000
BX897700_07	700001	810000
BX897700_08	800001	910000
BX897700_09	900001	1010000
BX897700_10	1000001	1110000
BX897700_11	1100001	1210000
BX897700_12	1200001	1310000
BX897700_13	1300001	1410000
BX897700_14	1400001	1510000
BX897700_15	1500001	1581384

Continuation (2 of 16) of BX897700 from base 100001 (BX897700 Bartonella quintana str. Tc

Query Match 63.9%; Score 21.2; DB 15; Length 110000;
Best Local Similarity 54.8%; Pred.No.3.7e+02;
Matches 23; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42
|||:| ||| ||| |||:| ||| |||:| ||| |||:
Db 101554 GAGCAACCATTCGCTGACAAATATAAGATTACGATGTAAG 101595

RESULT 41
BX897700_02
WPCOMMENT
Sequence split into 16 fragments LOCUS BX897700 Accession BX897700

Fragment Name	Begin	End
BX897700_00	1	110000
BX897700_01	100001	210000
BX897700_02	200001	310000
BX897700_03	300001	410000
BX897700_04	400001	510000
BX897700_05	500001	610000
BX897700_06	600001	710000
BX897700_07	700001	810000
BX897700_08	800001	910000
BX897700_09	900001	1010000
BX897700_10	1000001	1110000
BX897700_11	1100001	1210000

```

BX897700.12 1200001 1310000
BX897700.13 1300001 1410000
BX897700.14 1400001 1510000
BX897700.15 1500001 1581384
Continuation (3 of 16) of BX897700 from base 200001 (BX897700 Bartonella quintana str. T

Query Match 63.9%; Score 21.2; DB 15; Length 110000;
Best Local Similarity 54.8%; Pred. No. 3.7e+02;
Matches 23; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GCNCAVCGTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
Db 1554 GAGCAACCATTCGCTGAACAATATAAAGATTACGATGAAG 1595

RESULT 42
AC069465 AC069465 203946 bp DNA linear HTG 06-SEP-2000
DEFINITION Mus musculus clone RP23-23L3, WORKING DRAFT SEQUENCE, 18 unordered
pieces.
AC069465 AC069465.2 GI:9972306
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 203946)
AUTHORS McCombie,W.R., Baker,J.P., Bahret,A., Bal,H., Dedhia,N.N., de la
Bastide,M., Huang,E.N., King,L., Kirchoff,K.A., Miller,B.,
Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R., Rodriguez,M.A.,
Shah,R.S., Shether,M., Spiegel,L.A., Toth,K. and Vil,M.D.
TITLE Mouse Genomic Sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 203946)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
COMMENT On Sep 6, 2000 this sequence version replaced gi:8134853.
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: http://www.cshl.org/genseq
Contact: mcombie@cshl.org
----- Project Information
Center project name: RP23-23L3
Center clone name: RP23-23L3
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 30687: contig of 30687 bp in length
* 30688: gap of unknown length
* 30855: contig of 27293 bp in length
* 58148: gap of unknown length
* 58314: gap of unknown length
* 82613: contig of 24299 bp in length
* 82614: gap of unknown length
* 82780: gap of unknown length
* 106533: contig of 23753 bp in length
* 106534: gap of unknown length
* 106700: gap of unknown length
* 106701: contig of 16755 bp in length
* 123456: gap of unknown length
* 123623: contig of 13904 bp in length
* 137527: gap of unknown length
* 137693: contig of 9910 bp in length

```

```

* 147604 147770: gap of unknown length
* 147771 156659: contig of 8889 bp in length
* 156826: gap of unknown length
* 156827 156874: contig of 8848 bp in length
* 156875 15841: gap of unknown length
* 165842 174421: contig of 8580 bp in length
* 174422 174588: gap of unknown length
* 174589 181828: contig of 7240 bp in length
* 181829 181995: gap of unknown length
* 181996 187383: contig of 5388 bp in length
* 187384 187550: gap of unknown length
* 187551 192285: contig of 4735 bp in length
* 192286 192452: gap of unknown length
* 192453 196361: contig of 3909 bp in length
* 196362 196528: gap of unknown length
* 196529 199523: contig of 2995 bp in length
* 199524 199690: gap of unknown length
* 199691 201663: contig of 1972 bp in length
* 201663 201829: gap of unknown length
* 201830 203544: contig of 1715 bp in length
* 203545 203710: gap of unknown length
* 203711 203946: contig of 236 bp in length.
FEATURES
source Location/Qualifiers
1. 203946
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-23L3"
30688..30854
/estimated_length=unknown
58148..58314
/estimated_length=unknown
82614..82780
/estimated_length=unknown
106534..106700
/estimated_length=unknown
123456..123622
/estimated_length=unknown
137527..137693
/estimated_length=unknown
147604..147770
/estimated_length=unknown
156660..156826
/estimated_length=unknown
165675..165841
/estimated_length=unknown
174422..174588
/estimated_length=unknown
181829..181995
/estimated_length=unknown
187384..187550
/estimated_length=unknown
192286..192452
/estimated_length=unknown
196362..196528
/estimated_length=unknown
199524..199690
/estimated_length=unknown
201663..201829
/estimated_length=unknown
203545..203710
/estimated_length=unknown
ORIGIN
Query Match 63.9%; Score 21.2; DB 12; Length 203946;
Best Local Similarity 58.8%; Pred. No. 4.2e+02;
Matches 20; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
Qy 4 CAYCGTGGCCNTCCNTAYARGAYTAYGARG 37
Db 163807 CACACCTGGCCATCACAGTACAAAACATATCAGG 163840

```


JOURNAL

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jun 29, 2005 this sequence version replaced gi:62420177. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FELLY

Center clone name: CH240-70K12

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 97200 bases at least Q40

Consensus quality: 97183 bases at least Q30

Consensus quality: 97578 bases at least Q20

Estimated insert size: 99554; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 4537: contig of 4537 bp in length
* 4538 4587: gap of 50 bp
* 4588 37386: contig of 32799 bp in length
* 37387 37436: gap of 50 bp
* 37437 84977: contig of 47541 bp in length
* 84978 85027: gap of 50 bp
* 85028 96411: contig of 11384 bp in length
* 96412 96680: gap of 269 bp
* 96681 98142: contig of 1462 bp in length.

FEATURES
source

1. 98142
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-70K12"

4538. 4587
/estimated_length=50

37387. 37436
/estimated_length=50

84978. 85027
/estimated_length=50

96412. 96680
/estimated_length=269

ORIGIN

Query Match 63.3% Score 21; DB 12; Length 98142;

Best Local Similarity 58.3%; Pred. No. 4.6e+02;

Matches 21; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 6 YGCNTGGCCNTCCNTAYAAAGATYATGARGTNA 41

::: ||||| ||||| : ||| : ||| : |||

Db 81013 TGCATGGCCATCACCACCATGATGATAATGAATAA 80978

RESULT 45

AC066598

LOCUS

DEFINITION Homo sapiens chromosome 3 clone RP11-542K24 map 3p, complete sequence.

AC066598 142190 bp DNA linear

AC066598.5 GI:10800219

AC066598

HTG

KEYWORDS

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 142190)

Niu, Y., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,

Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,

He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G.,

Li, J., Li, L., Li, S., Li, T., Liu, X., Liu, N., Liu, B., Liu, Y., Li, W.,

Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M.,

Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, L.,

Wang, L., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F.,

Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G.,

Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X.,

Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Unpublished

Chromosome 3p genomic sequence

2 (bases 1 to 142190)

Wang, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B.,

Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y.,

Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y.,

Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,

Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.

and Yang, H.

Direct Submission

Submitted (25-APR-2000) Human Genomic Center, Institute of

Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing

100101, P.R.China

3 (bases 1 to 142190)

Niu, Y., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,

Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,

He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G.,

Li, J., Li, L., Li, S., Li, T., Liu, X., Liu, N., Liu, B., Liu, Y., Li, W.,

Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M.,

Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, L.,

Wang, L., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F.,

Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G.,

Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X.,

Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Direct Submission

Submitted (13-OCT-2000) Human Genomic Center, Institute of

Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing

100101, P.R.China

On Oct 13, 2000 this sequence version replaced gi:10719850.

-----Genome Center

Center: Beijing Center

Center code: Beijing

Website: <http://hgsc.igtp.ac.cn>

<http://www.genomics.org.cn>

Contact: hgsc@igtp.ac.cn

----- Project Information

Center project name: 1% project

Center clone name: RP11-542K24

----- Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator; ET 55% of reads

Chemistry: Dye-terminator Big Dye; 45% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 14432 bases at least Q40

Consensus quality: 144770 bases at least Q30

Consensus quality: 144982 bases at least Q20

REFERENCE
AUTHORS

116361..116460
/estimated_length=unknown
123050..123149
/estimated_length=unknown
126646..126745
/estimated_length=unknown
131953..132052
/estimated_length=unknown
135712..135811
/estimated_length=unknown
140617..140716
/estimated_length=unknown
143564..143663
/estimated_length=unknown
145190..145289
/estimated_length=unknown
148745..148844
/estimated_length=unknown
147966..148065
/estimated_length=unknown
149319..149418
/estimated_length=unknown
150672..150771
/estimated_length=unknown

ORIGIN

Query Match 63.3% Score 21; DB 12; Length 151830;
Best Local Similarity 51.2%; Pred. No. 5e+02;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 2 CNCAYGCGTGGCCNCCNTCCATGAAAGACGATGAGTTAA 42

Db 39728 CTCTCTCAGCCTTCCCATGAAAGACGATGAGTTAA 39768

RESULT 47

AC024164/c
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-350A17 map 3p, complete
sequence.

AC024164

AC024164.6 GI:13899382

HTG.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 152580)

Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z.,
He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
Li,W., Li,X., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 152580)

Li,W., Hu,S., Dong,W., Zhang,X., Wang,J., Zhang,Y., Zhang,H.,
Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y.,
Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H.,
Liu,Y., Li,G., Li,C., Li,C., Bao,Q., Bao,J., Song,L., Zhang,L.,
Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L.,
Feng,X., Yu,J. and Yang,H.

Direct Submission

Submitted (25-FEB-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

FEATURES
source

1..152580
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

3 (bases 1 to 152580)

Ning,Y., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,W.,
Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W.,
Li,W., Li,X., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.

Direct Submission

Submitted (13-OCT-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

4 (bases 1 to 152580)

Niu,Y., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
Liu,Y., Li,W., Li,X., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.

Direct Submission

Submitted (02-FEB-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

5 (bases 1 to 152580)

Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z.,
He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
Li,W., Li,X., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.

Direct Submission

Submitted (01-MAY-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

On May 1, 2001 this sequence version replaced gi:12656790.

-----Genome Center

Center:Beijing Center

Center code:Beijing

Website:http://hgci.gtp.ac.cn

http://www.genomics.org.cn

Contact:hgci@gtp.ac.cn

----- Project Information

Center project name:1# project

Center clone name: RP11-350A17

----- Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator: ET 55% of reads

Chemistry: Dye-terminator: Big Dye; 45% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 154635 bases at least Q40

Consensus quality: 155213 bases at least Q30

Consensus quality: 155499 bases at least Q20

Insert size: 152580; sum-of-contigs

Quality coverage: 9.66x in Q20 bases;sum-of-contigs

----- Location/Qualifiers

1..152580

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"


```
/chromosome="3"
/map="3p"
/clone="RP11-350A17"

Query Match      63.3%; Score 21; DB 5; Length 152580;
Best Local Similarity 51.2%; Pred. No. 5e+02;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 2 CMCAYCCTGGCCCTCCNCCNTAAYAGCATYAGGTGTAAR 42
Db 151951 CTTCTCTCAGCTTCCCATGTGAAGACGATGAGGTAAA 151911

RESULT 48
AC024219
LOCUS          154010 bp      DNA      linear      HTG 06-SEP-2000
DEFINITION    Homo sapiens chromosome 3 clone RP11-350A17, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
AC024219
AC024219.11 GI:9438327
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE    1 (bases 1 to 154010)
AUTHORS     Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hoques,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovari,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucter,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,S., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shan,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugeng,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wabba,M.,
            Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D., and
            Gibbs,R.
            Direct Submission
            Unpublished
            2 (bases 1 to 154010)
            Worley,K.C.
            Direct Submission
            Submitted (28-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Jul 25, 2000 this sequence version replaced gi:8699729.
            ----- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ----- Project Information
            Center project name: HAIE
            Center clone name: RP11-350A17
            ----- Summary Statistics
            Sequencing vector: M13; L08821
            Chemistry: Dye-terminator Big Dye. 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 140016 bases at least Q40
            Consensus quality: 145094 bases at least Q30
            Consensus quality: 147945 bases at least Q20
            Estimated insert size: 150053; sum-of-contigs estimation
            Quality coverage: 0x in Q20 bases; agarose-gel estimation

-----
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 13 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  * 1 28396: contig of 28396 bp in length
  * 28397 28496: gap of unknown length
  * 58151: contig of 29655 bp in length
  * 58152 58251: gap of unknown length
  * 58252 77456: contig of 19205 bp in length
  * 77457 77556: gap of unknown length
  * 77557 94589: contig of 17033 bp in length
  * 94590 94689: gap of unknown length
  * 94690 108289: contig of 13600 bp in length
  * 108290 108389: gap of unknown length
  * 108390 119719: contig of 11330 bp in length
  * 119720 119819: gap of unknown length
  * 119820 128258: contig of 8439 bp in length
  * 128259 128358: gap of unknown length
  * 128359 135273: contig of 6915 bp in length
  * 135274 135373: gap of unknown length
  * 135374 141772: contig of 6399 bp in length
  * 141773 147268: contig of 5396 bp in length
  * 147269 147368: gap of unknown length
  * 147369 149978: contig of 2610 bp in length
  * 149979 150078: gap of unknown length
  * 150079 152110: contig of 2032 bp in length
  * 152111 152210: gap of unknown length
  * 152211 154010: contig of 1800 bp in length.
            Location/Qualifiers
            1. 154010
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="3"
            /clone="RP11-350A17"
            /estimated_length=unknown
            58152..58251
            /estimated_length=unknown
            77457..77556
            /estimated_length=unknown
            94590..94689
            /estimated_length=unknown
            108290..108389
            /estimated_length=unknown
            119720..119819
            /estimated_length=unknown
            128259..128358
            /estimated_length=unknown
            135274..135373
            /estimated_length=unknown
            141773..141872
            /estimated_length=unknown
            147269..147368
            /estimated_length=unknown
            149979..150078
            /estimated_length=unknown
            152111..152210
            /estimated_length=unknown

ORIGIN
Query Match      63.3%; Score 21; DB 12; Length 154010;
Best Local Similarity 51.2%; Pred. No. 5e+02;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;
            -----
            Query Match      63.3%; Score 21; DB 12; Length 154010;
            Best Local Similarity 51.2%; Pred. No. 5e+02;
            Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;
```


JOURNAL

Submitted (31-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 31, 2001 this sequence version replaced gi:15029434.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source

QUALSTAT-REPORT.

Location/Qualifiers
1. .178752
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="Rp11-554F8"

misc_feature

complement(1..10232)
/notes="overlaps bases 10232..1 of clone AC055710"
/function="clone overlap"

672..705

/rpt_family="(TG)n"

repeat_region

1233..1253

/rpt_family="AT_rich"

repeat_region

1698..1894

/rpt_family="AluSq"

repeat_region

2952..3265

/rpt_family="L1PA13"

repeat_region

3356..3386

/rpt_family="AT_rich"

repeat_region

3562..3703

/rpt_family="CT-rich"

repeat_region

4704..4734

/rpt_family="AT_rich"

repeat_region

complement(4956..5141)

/rpt_family="MIR"

STS

6568..6869

/standard_name="174894"

repeat_region

7016..7065

/rpt_family="A-rich"

repeat_region

complement(10872..10947)

repeat_region

/rpt_family="MER116"

repeat_region

complement(11932..12112)

repeat_region

/rpt_family="MIR"

repeat_region

complement(12199..12263)

repeat_region

/rpt_family="MIR"

repeat_region

14167..14536

/rpt_family="MER7A"

repeat_region

14709..14736

/rpt_family="L2"

repeat_region

14737..14788

/rpt_family="MIR"

repeat_region

complement(15365..15720)

repeat_region

/rpt_family="THE1C"

repeat_region

16309..16580

/rpt_family="L1PA8"

repeat_region

16587..16610

/rpt_family="(TA)n"

repeat_region

complement(17720..17939)

repeat_region

/rpt_family="MIR"

repeat_region

complement(18069..18242)

repeat_region

/rpt_family="MER104"

repeat_region

18803..18957

/rpt_family="MIR"

repeat_region

19158..19196

/rpt_family="AT_rich"

repeat_region

20075..20095

/rpt_family="AT_rich"

repeat_region

20307..20452

/rpt_family="MIR"

repeat_region

complement(20748..21216)

repeat_region

/rpt_family="L1MA3"

repeat_region

23739..23768

/rpt_family="(TAAA)n"

repeat_region

24295..24324

/rpt_family="AT_rich"

repeat_region

complement(26072..26605)

repeat_region

/rpt_family="L2"

repeat_region

26814..26964

/rpt_family="MER5A"

repeat_region

27396..27424

/rpt_family="AT_rich"

repeat_region

27759..27849

/rpt_family="MIR"

repeat_region

complement(28789..28980)

repeat_region

/rpt_family="MIR"

repeat_region

complement(29219..29266)

repeat_region

/rpt_family="MER5A"

repeat_region

29289..29323

/rpt_family="MER5A"

repeat_region

complement(29367..29774)

repeat_region

/rpt_family="MSTA"

repeat_region

30098..30119

/rpt_family="AT_rich"

repeat_region

30147..30274

/rpt_family="MIR"

repeat_region

complement(30290..30382)

repeat_region

/rpt_family="MER91C"

repeat_region

30819..30842

/rpt_family="AT_rich"

repeat_region

complement(30843..31152)

repeat_region

/rpt_family="AluY"

repeat_region

31153..31174

Query Match 63.3%; Score 21; DB 5; Length 178752;

Best Local Similarity 51.2%; Pred.No 5.1e+02;

Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 2 CMCAGCCTGGCCGTCNCCTTCAATGAAGATTATGAGGACAAA 97917

DB 97957 CTCTTCTTCTCTCTTCAATGAAGATTATGAGGACAAA 97917

RESULT 53	AC087589	181736 bp	DNA	linear	PRI 08-NOV-2002
AC087589/c					
LOCUS	Homo sapiens chromosome 3 clone RP11-1102N22 map 3p, complete				
DEFINITION	sequence.				
ACCESSION	AC087589	3	GI:24270662		
VERSION	HTG.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 181736)				
AUTHORS	Niu, Y., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Song, S., Sun, M., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, L., Wang, J., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				
TITLE	Chromosome 3p genomic sequence				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 181736)				
AUTHORS	Niu, Y., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-JAN-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China				
REFERENCE	3 (bases 1 to 181736)				
AUTHORS	Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAY-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China				
REFERENCE	4 (bases 1 to 181736)				
AUTHORS	Niu, Y., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-OCT-2002) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China				
REFERENCE	5 (bases 1 to 181736)				
AUTHORS	Niu, Y., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-OCT-2002) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China				

Consensus quality: 196152 bases at least Q40
 Consensus quality: 196403 bases at least Q30
 Consensus quality: 196672 bases at least Q20
 Insert size: 197255; sum-of-contigs
 Insert size: 229395; 8.2% error; agarose-fp
 Quality coverage: 11.17x in Q20 bases; sum-of-contigs Quality
 Coverage: 9.69x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 30719: contig of 30719 bp in length
 * 30720 30819: gap of 100 bp
 * 30820 39556: contig of 8737 bp in length
 * 39557 39636: gap of 100 bp
 * 39637 78034: contig of 38378 bp in length
 * 78035 78134: gap of 100 bp
 * 78135 130267: contig of 52133 bp in length
 * 130268 130367: gap of 100 bp
 * 130368 147194: contig of 16827 bp in length
 * 147195 147294: gap of 100 bp
 * 147295 197755: contig of 50461 bp in length.

FEATURES

source

1. 197755
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="17"
 /clone="CH29-545C14"
 /clone_lib="CHORI-29"

misc_feature

1 30719
 /note="assembly_fragment:00416
 fragment_chain:1
 clone_end:T7
 vector_side:left"
 30820..39556
 /note="assembly_fragment:00013
 fragment_chain:1"

misc_feature

39657..78034
 /note="assembly_fragment:00942
 fragment_chain:1"
 78135..130267
 /note="assembly_fragment:02424
 fragment_chain:1"

misc_feature

130368..147194
 /note="assembly_fragment:00150
 fragment_chain:1"
 147295..197755
 /note="assembly_fragment:01582
 fragment_chain:1
 clone_end:SP6
 vector_side:right"

misc_feature

147295..197755
 /note="assembly_fragment:01582
 fragment_chain:1
 clone_end:SP6
 vector_side:right"

misc_feature

147295..197755
 /note="assembly_fragment:01582
 fragment_chain:1
 clone_end:SP6
 vector_side:right"

ORIGIN

Query Match 63.1%; Score 21; DB 12; Length 197755;
 Best Local Similarity 55.3%; Pred.No. 5.2e+02;
 Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAVGCNCGCCNCCNCAATAYARGATYAGTGT 38

Db 126545 GCTGCTCCCTGGCTTCTCCTACAGGTCTCTGGGT 126582

RESULT 57

AC172772

LOCUS

DEFINITION AC172772 215523 bp DNA linear HTG 09-FEB-2006
 Bos taurus clone CH240-267C23, *** SEQUENCING IN PROGRESS ***, 19
 unordered pieces.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AC172772 GI:87081468
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 Bos taurus (cattle)

ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.

REFERENCE
 AUTHORS

1 (bases 1 to 215523)
 Muzny, D., Adams, C., Agbai, II, O., Allen, C., Alsbrooks, S., Archer, P.,
 Arréondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
 Beraducci, A., Biswalo, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
 Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
 Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
 Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
 Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
 Anda, C., Delgado, O., Denson, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
 Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
 Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
 Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
 Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
 Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
 Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlein, K.,
 Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
 Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
 Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B.,
 Howell, L.T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
 Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
 Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,
 Leal, S., Lee, K., Lee, S., LeGall, F.I., Lemon, S., Lewis, L., Li, B.,
 Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
 London, P., Lopez, J., Lorensunhewa, L., Lozado, R., Luk, T., Madu, R.,
 Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
 McClelland, H., McPherson, J., Mercadado, C., Metzker, M.,
 Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
 Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,
 Nott, A., Nwaokeme, O., Obregon, M., Ochi-Okorie, C., Odeh, E.,
 Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
 Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
 Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiros, J.,
 Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
 Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W.,
 Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
 Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
 Svatok, A., Taylor, E., Taylor, T., Thomas, N., Thornt, R., Thornton, R.,
 Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villasana, D., Virk, D.,
 Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
 Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
 Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
 Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.A.

TITLE
 JOURNAL
 REFERENCE

2 (bases 1 to 215523)

AUTHORS
 JOURNAL

Worley, K.C.

TITLE
 JOURNAL

Submitted (23-NOV-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 AUTHORS
 CONSTM
 TITLE
 JOURNAL

3 (bases 1 to 215523)

REFERENCE
 AUTHORS
 CONSTM
 TITLE
 JOURNAL

Bovine Genome Sequencing Consortium

COMMENT

Submitted (09-FEB-2006) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Feb 9, 2006 this sequence version replaced gi:82654423.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence


```

REFERENCE
AUTHORS      Parkway, St. Louis, MO 63108, USA
TITLE         4 (bases 1 to 228141)
JOURNAL       Wilson, R.K.
              Direct Submission
              Submitted (05-MAR-2005) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE     5 (bases 1 to 228141)
AUTHORS      Wilson, R.K.
TITLE         Direct Submission
JOURNAL       Submitted (08-MAR-2005) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
COMMENT       On Mar 5, 2005 this sequence version replaced gi:57470749.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
              ----- Summary Statistics
              -----
              Center project name: M_BA0023B10
              -----

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
and Minako Tateni in the laboratory of Pieter de Jong
(http://www.chori.org) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (http://www.resgen.com) or
Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1. .228141
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="17"
   /map="17"
   /clone="RP23-23B10"
   /clone_lib="RPCI-23"
   91156..91170
misc_feature
   /note="Sequence derived from PCR product of project DNA."
91258..91361
misc_feature
   /note="Sequence derived from PCR product of project DNA."
91783..91793
   /note="Sequence derived from one plasmid subclone."
125291..125321
   /note="Sequence derived from one plasmid subclone."
130859..131057
   /note="Sequence derived from one plasmid subclone."
167634..167696
   /note="Sequence derived from one plasmid subclone."
189541..189685
   /note="Sequence derived from one plasmid subclone."
190009..190028
   /note="Sequence derived from one plasmid subclone."

REFERENCE
AUTHORS      190075..190144
TITLE         /note="Unresolved simple sequence repeat."
JOURNAL       190145..190162
              /note="Sequence derived from one plasmid subclone."
              199991..200047
              /note="Sequence derived from one plasmid subclone."
ORIGIN
Query Match      63.3%; Score 21; DB 6; Length 228141;
Best Local Similarity 55.3%; Pred. NO. 5.4e+02;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY      1 GCNCAAGCTGGCCATCNCNTAYARGAYTAYGARGT 38
        |||:|||||:|||||:|||||:|||||:|||||:
DB      95053 GCTGCTCCTGGCTTCTCCCTACAGGTCCTGCGGT 95016

RESULT 59
AC095735
LOCUS      Rattus norvegicus clone CH230-9A22, *** SEQUENCING IN PROGRESS ***,
DEFINITION 2 unordered pieces.
ACCESSION AC095735
VERSION    AC095735.5 GI:22773198
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridea; Muridae; Murinae; Rattus;
            1 (bases 1 to 229726)
REFERENCE  Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
            Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
            Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H.,
            Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
            Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
            Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
            Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
            Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
            Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
            Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
            Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
            Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
            Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
            Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
            Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
            Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
            Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
            Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
            Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
            Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
            Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
            Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
            Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
            Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
            Lorensuhewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,
            Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
            Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
            Manthey, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
            Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
            Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, J., Nair, L.,
            Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
            Nwaokemeleh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
            Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
            Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
            Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
            Reilly, B., Reilly, M., Ren, Y., Reuter, M., Rose, R., Ruiz, S.J.,
            Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
            Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
            Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajd, D.,
            Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
            Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
            Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

```


LOCUS	AC109060	246535 bp	DNA	linear	HTG 08-OCT-2002
DEFINITION	Rattus norvegicus clone CH230-9506, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.				
ACCESSION	AC109060				
VERSION	AC109060.4 GI:23195511				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 246535)				
AUTHORS	Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, I., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Hawlak, P., Haves, A., Henderson, N., Hernandez, J., Herrandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lepow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steelme, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K., Valdes, R., Vera, V., Villabana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
AUTHORS	2 (bases 1 to 246535)				
JOURNAL	Worley, K.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 246535)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (08-OCT-2002) Human Genome Sequencing Center, Department				

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 19, 2002 this sequence version replaced gi:21737852. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPVU
Center clone name: CH230-9506
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 194634 bases at least Q40
Consensus quality: 199097 bases at least Q30
Consensus quality: 202096 bases at least Q20
Estimated insert size: 216622; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 90577: contig of 90577 bp in length
* 90578 90677: gap of unknown length
* 90678 244089: contig of 153412 bp in length
* 244090 244189: gap of unknown length
* 244190 246535: contig of 2346 bp in length.

FEATURES
Location/Qualifiers
1..246535
/organism="Rattus norvegicus"
/mol_type="Genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9506"
misc_feature 1..1367
/note="wgs contig"
misc_feature 4513..5520
/note="wgs contig"
misc_feature 26837..27898
/note="wgs contig"
gap 90578..90677
/estimated_length=unknown
misc_feature 17840..179232
/note="wgs contig"
misc_feature 179283..182690
/note="wgs contig"
gap 244090..244189
/estimated_length=unknown

ORIGIN
Query Match 63.3%; Score 21; DB 12; Length 246535;
Best Local Similarity 58.3%; Pred. No. 5;Se+02;
Matches 21; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 5 AYCNTGGCCNTCCNTAYARGAYTAYGARTNA 40


```
Best Local Similarity 55.0%; Pred. No. 5.9e+02;
Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCCTGCCNCCNTAYARGAYTAYGARGTNA 40
    |||||
Db 58698 GCCCACCACCTGTCCCTCCAGTAGAAGGACTGCCGAGCGCA 58659

RESULT 66
CR354544
LOCUS
DEFINITION
Danio rerio chromosome 23 clone CH211-273A21, WORKING DRAFT
SEQUENCE, 3 unordered pieces.
ACCESSION
CR354544
VERSION
CR354544.4 GI:74197629
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 139509)
McLaren, S.
Direct Submission
Submitted (01-SEP-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
On Sep 6, 2005 this sequence version replaced GI:45581022.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zk23G9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 138778 bases at least Q40
Consensus quality: 138874 bases at least Q30
Consensus quality: 138966 bases at least Q20
Insert size: 139309; sum-of-contigs
Insert size: 166839; 1.5% error; agarose-ff
Quality coverage: 8.53x in Q20 bases; sum-of-contigs Quality
coverage: 7.17x in Q20 bases; agarose-ff
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 12730: contig of 12730 bp in length
* 12731 12830: gap of 100 bp
* 12831 43204: contig of 30374 bp in length
* 43205 43304: gap of 100 bp
* 43305 139509: contig of 96205 bp in length.
Location/Qualifiers
1. 139509
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="23"
/clone="CH211-273A21"
/clone_lib="CHORI-211"
1. 12730
/notes="assembly fragment:01346
fragment_chain:1
clone_end:SP6
vector_side:left"

misc_feature
12831..43204
/notes="assembly fragment:01493
fragment_chain:1"
43305..139509
/notes="assembly fragment:00001
fragment_chain:1"

ORIGIN
Query Match 62.7%; Score 20.8; DB 12; Length 139509;
Best Local Similarity 52.4%; Pred. No. 6.2e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCCTGCCNCCNTAYARGAYTAYGARGTNAAR 42
    |||||
Db 55915 GCGTGTGCAAGCCCTCTTCTTCAAGCCCTATGATGTGAAA 55956

RESULT 67
AL662881
LOCUS
DEFINITION
Mouse DNA sequence from clone RP23-321C9 on chromosome 11 Contains
part of the gene for a novel protein similar to Tensin Tns,
complete sequence.
ACCESSION
AL662881
VERSION
AL662881.10 GI:20068665
KEYWORDS
HTG; Tensin.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 154959)
White, S.
Direct Submission
Submitted (04-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
vegas@sanger.ac.uk
Clone requests:
clonerequest@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced GI:19572489.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
-----
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
RP23-321C9 is from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.
Location/Qualifiers
1. 154959
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-321C9"
/clone_lib="RPCI-23"
join(complement(AL645636.22:38537..38593),
```



```

KEYWORDS
SOURCE      Xenopus tropicalis (Silurana tropicalis)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 165987)
AUTHORS    DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE      Direct Submission
JOURNAL
REFERENCE   2 (bases 1 to 165987)
AUTHORS    DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL    Submitted (18-JAN-2004) Production Genomics Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
            94598-1698, USA
REFERENCE   3 (bases 1 to 165987)
AUTHORS    DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE      Direct Submission
JOURNAL    Submitted (13-APR-2004) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT    On Apr 13, 2004 this sequence version replaced gi:41016131.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www.shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source           1..165987
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="CH216-117C13"
ORIGIN
Query Match      62.7%; Score 20.8; DB 11; Length 165987;
Best Local Similarity 52.4%; Pred. No. 6.4e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY   1  GCNAYGCGTGGCCNTCNCNTAYARGAVTAYGAGCTNAAR 42
      |||||
DB   30040  GACCAAGCTTGCTGCTCCAGACAAAGTTTATCAGCTGAAA 29999
      |||||

RESULT 69
AC131779
LOCUS      AC131779          171978 bp    DNA    linear    ROD 28-JAN-2005
DEFINITION Mus musculus BAC clone RP24-336D11 from 8, complete sequence.
ACCESSION  AC131779
VERSION    AC131779.4  GI:56236242
KEYWORDS  HTG.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 171978)
AUTHORS    Tomlinson, C., Cotton, M. and Haakenson, W.
TITLE      The sequence of Mus musculus BAC clone RP24-336D11
JOURNAL    Unpublished (2001)
REFERENCE   2 (bases 1 to 171978)
AUTHORS    McPherson, J.D. and Waterston, R.H.
TITLE      Direct Submission
JOURNAL    Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3 (bases 1 to 171978)
AUTHORS    Wilson, R.K.
TITLE      Direct Submission
JOURNAL    Submitted (28-SEP-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4 (bases 1 to 171978)
AUTHORS    Wilson, R.K.
TITLE      Direct Submission
JOURNAL    Submitted (28-SEP-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5 (bases 1 to 171978)
AUTHORS    Wilson, R.K.
TITLE      Direct Submission
JOURNAL    Submitted (03-DEC-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   6 (bases 1 to 171978)
AUTHORS    Wilson, R.K.
TITLE      Direct Submission
JOURNAL    Submitted (28-JAN-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT    On Dec 3, 2004 this sequence version replaced gi:52782655.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@wustl.edu
            ----- Summary Statistics
            -----
            Center project name: M_BB0336D11
            -----
NOTICE
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu
SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org
NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source           1..171978
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="8"
                     /map="8"
                     /clone="RP24-336D11"
                     /clone_lib="RPCI-24"
                     /notes="Sequence derived from PCR product of project DNA."
     misc_feature     117242..117379
                     /notes="Unresolved bases."
     unsure           117361
                     /notes="Unresolved bases."
     unsure           117376
                     /notes="Unresolved bases."
     repeat_region    129522..129523
                     /transposon="Bacterial transposon insertion in clone
                     excised here."
     unsure           137687..138228
                     /notes="Unresolved simple sequence repeat."
     misc_feature     138229..138312
                     /notes="Sequence derived from PCR product of project DNA."
     unsure           140133..140250
                     /note="Unresolved simple sequence repeat."
ORIGIN
Query Match      62.7%; Score 20.8; DB 6; Length 171978;
Best Local Similarity 55.0%; Pred. No. 6.4e+02;
Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

```



```

misc_feature 7419. 22998
/note="assembly_fragment"
22999. 23098
/estimated_length=unknown
misc_feature 23099. 35627
/note="assembly_fragment"
35628. 35727
/estimated_length=unknown
misc_feature 35728. 55335
/note="assembly_fragment"
55336. 55435
/estimated_length=unknown
misc_feature 55436. 77882
/note="assembly_fragment"
77883. 77982
/estimated_length=unknown
misc_feature 77983. 101387
/note="assembly_fragment"
101388. 101487
/estimated_length=unknown
misc_feature 101488. 138406
/note="assembly_fragment"
138407. 138506
/estimated_length=unknown
misc_feature 138507. 162815
/note="assembly_fragment"
162816. 162915
/estimated_length=unknown
misc_feature 162916. 174929
/note="assembly_fragment
clone_end:T7
vector_side:right"

ORIGIN
Query Match 62.7%; Score 20.8; DB 12; Length 174929;
Best Local Similarity 55.0%; Pred. No. 6.4e+02;
Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 GCNCAVCNTGGCCNTCCNTAYAGATYAGATGNA 40
Db 150532 GCATGGCAGGGCTCGCCATATTTGGACAATGATTTA 150493

RESULT 71
AC132640
LOCUS Rattus norvegicus clone CH230-339B14, WORKING DRAFT SEQUENCE.
DEFINITION AC132640
AC132640
VERSION AC132640.3 GI:25007393
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (bases 1 to 194709)
Muzny, D., Maré, M., Metzker, M., Lee, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesat, H., Chen, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devila, M., Davis, C., Davy-Carrroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

```

```

TITLE JOURNAL
REFERENCE 2 (bases 1 to 194709)
Rat Genome Sequencing Consortium.
Direct Submission
Direct Submission
Submitted (02-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 194709)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23908481.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBQF
Center clone name: CH230-339B14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 188047 bases at least Q40
Consensus quality: 189447 bases at least Q30
Consensus quality: 190412 bases at least Q20
Estimated insert size: 195068; sum-of-contigs estimation

```

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loulseged, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, W., Savery, G., Scherer, S., Scott, G., Shateman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 194709)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (02-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 194709)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23908481.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KBQF

Center clone name: CH230-339B14

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 188047 bases at least Q40

Consensus quality: 189447 bases at least Q30

Consensus quality: 190412 bases at least Q20

Estimated insert size: 195068; sum-of-contigs estimation


```

-----
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces
  is believed to be correct as given, however the sizes
  of the gaps between them are based on estimates that have
  provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 194709: contig of 194709 bp in length.
Location/Qualifiers
1. 194709
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-339B14"
1. 3065
/notes="wgs end _extension
clone_end:T7"
7050..7929
/notes="clone_boundary
clone_end:T7"
misc_feature
1. 3065
/notes="wgs end _extension
clone_end:T7"
7050..7929
/notes="clone_boundary
clone_end:Sp6"
end_sequence:B2164293"
misc_feature
1. 194709: contig of 194709 bp in length.
end_sequence:B2164293"
ORIGIN
Query Match 62.7%; Score 20.8; DB 12; Length 194709;
Best Local Similarity 52.4%; Pred. No. 6.6e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNACGTCNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169254 GCCAAGCTGTTGGCCAGTCACAGACAGAGAGTAAAG 169295
|||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 72
AC126211 195892 bp DNA linear HTG 12-OCT-2002
LOCUS Rattus norvegicus clone CH230-41316, WORKING DRAFT SEQUENCE.
DEFINITION AC126211
ACCESSION AC126211.3 GI:23907800
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Bliswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

```

```

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.I., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smaiz,D.,
Speed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 195892)
Worley,K.C.
Direct Submission
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 195892)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (12-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 12, 2002 this sequence version replaced gi:21700469.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYFP
Center clone name: CH230-41316
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 171920 bases at least Q40
Consensus quality: 173922 bases at least Q30
Consensus quality: 175324 bases at least Q20

```



```

Estimated insert size: 176162; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces
  is believed to be correct as given, however the sizes
  of the gaps between them are based on estimates that have
  been provided by the submitter.
* This sequence will be replaced
  as soon as it is available and
  as soon as it is preserved.
* The accession number will be preserved.
* 1 195892: contig of 195892 bp in length.
FEATURES             Location/Qualifiers
     source            1..195892
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-41316"
     misc_feature      1..1313
                        /note="wgs_end_extension"
                        clone_end:T7"
                        1815..2881
                        /note="wgs_end_extension"
                        clone_end:T7"
     misc_feature      complement(11158..12220)
                        /note="clone_boundary"
                        clone_end:T7"
                        site:Mbol
                        end_sequence:RXBER51TJ"
     misc_feature      194721..195289
                        /note="clone_boundary"
                        clone_end:Sp6
                        site:Mbol
                        end_sequence:RXBER51TV"
ORIGIN
Query Match          62.7%; Score 20.8; DB 12; Length 195892;
Best Local Similarity 59.5%; Pred. No. 6.6e+02;
Matches 22; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCNCAVCGTGGCCNCCNCTAYAAAGAYTAYGARG 37
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109270 GTGCATTGTGGCCCTTCTTACAAAGATAGGAG 109306

RESULT 73
CT573319/c
LOCUS               210916 bp DNA linear HTG 30-JAN-2006
DEFINITION          Danio rerio chromosome 23 clone DKEY-3K23, *** SEQUENCING IN
                    PROGRESS ***, 24 unordered pieces.
ACCESSION            CT573319
VERSION              CT573319.2 GI:86197632
KEYWORDS             HTG; HTGS PHASE1.
SOURCE               Danio rerio (zebrafish)
ORGANISM             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                    Cypriniformes; Cyprinidae; Danio.
REFERENCE            1 (bases 1 to 210916)
AUTHORS              Sims,S.
TITLE               Direct Submission
JOURNAL             Submitted (29-JAN-2006) Wellcome Trust Sanger Institute, Hinxton,
                    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                    zfish-help@sanger.ac.uk Clone requests:
                    http://www.sanger.ac.uk/Projects/D\_rerio/faq/shtml#dataeight
COMMENT              On Jan 30, 2006 this sequence version replaced gi:85857250.
                    ----- Genome Center
                    Center: Wellcome Trust Sanger Institute
                    Center code: SC
                    Web site: http://www.sanger.ac.uk

```

```

Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK3K23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 198408 bases at least Q40
Consensus quality: 200547 bases at least Q30
Consensus quality: 202510 bases at least Q20
Insert size: 208616; sum-of-contigs
Insert size: 194803; 3.0% error; agarose-fp
Quality coverage: 3.97x in Q20 bases; sum-of-contigs Quality
coverage: 4.29x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
  consists of 24 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
  * 1 5453: contig of 5453 bp in length
  * 5454: gap of 100 bp
  * 5554: gap of 3825 bp in length
  * 5554: gap of 100 bp
  * 9379: contig of 10305 bp in length
  * 19783: gap of 100 bp
  * 19784: gap of 100 bp
  * 19884: contig of 19226 bp in length
  * 39109: gap of 100 bp
  * 39110: contig of 23194 bp in length
  * 62403: gap of 100 bp
  * 62404: gap of 100 bp
  * 62503: gap of 100 bp
  * 62504: gap of 11289 bp in length
  * 73793: gap of 100 bp
  * 73893: gap of 4983 bp in length
  * 78875: gap of 100 bp
  * 78976: gap of 8663 bp in length
  * 87639: gap of 100 bp
  * 87738: gap of 5389 bp in length
  * 87739: gap of 100 bp
  * 93128: gap of 100 bp
  * 93227: gap of 2021 bp in length
  * 93228: gap of 100 bp
  * 95348: gap of 100 bp
  * 95349: gap of 4536 bp in length
  * 99884: gap of 100 bp
  * 99885: gap of 4735 bp in length
  * 104719: gap of 100 bp
  * 104720: gap of 100 bp
  * 116507: contig of 11687 bp in length
  * 116507: gap of 100 bp
  * 116607: contig of 14034 bp in length
  * 130641: gap of 100 bp
  * 130741: gap of 8770 bp in length
  * 139510: gap of 100 bp
  * 139511: gap of 4788 bp in length
  * 139611: gap of 100 bp
  * 144399: gap of 2152 bp in length
  * 144999: gap of 100 bp
  * 146651: gap of 100 bp
  * 146751: contig of 3231 bp in length
  * 150082: gap of 100 bp
  * 150082: gap of 8113 bp in length
  * 158195: gap of 100 bp
  * 158295: gap of 6953 bp in length
  * 158295: gap of 100 bp
  * 165248: gap of 100 bp
  * 165348: contig of 10632 bp in length
  * 175980: gap of 100 bp
  * 176080: contig of 3605 bp in length
  * 179685: gap of 100 bp
  * 179785: contig of 5765 bp in length
  * 185550: gap of 100 bp
  * 185650: contig of 25267 bp in length.
FEATURES             Location/Qualifiers
     source            1..210916
                        /organism="Danio rerio"
                        /mol_type="genomic DNA"

```

```
/db_xref="taxon:7955"
/chromosomes="23"
/clone="DKEY-3K23"
/clone_lib="DanioKey"
1. .5453
/notes="assembly fragment:00258
fragment_chain:1"
misc_feature
5554. .9378
/notes="assembly fragment:00056
fragment_chain:1"
misc_feature
9479. .19783
/notes="assembly fragment:00422
fragment_chain:1"
misc_feature
19884. .39109
/notes="assembly fragment:00771
fragment_chain:1"
misc_feature
39210. .52403
/notes="assembly fragment:00877
fragment_chain:1"
misc_feature
62504. .73792
/notes="assembly fragment:00605
fragment_chain:1"
misc_feature
73993. .78875
/notes="assembly fragment:00167
fragment_chain:1"
misc_feature
78976. .87638
/notes="assembly fragment:00371
fragment_chain:2"
misc_feature
87739. .93127
/notes="assembly fragment:00195
fragment_chain:2"
misc_feature
93228. .95248
/notes="assembly fragment:00010
fragment_chain:2"
misc_feature
95349. .99884
/notes="assembly fragment:00094
fragment_chain:2"
misc_feature
99985. .104719
/notes="assembly fragment:00075
fragment_chain:2"
misc_feature
104820. .116506
/notes="assembly fragment:00538
fragment_chain:2"
misc_feature
116607. .130640
/notes="assembly fragment:00678
fragment_chain:2"
misc_feature
130741. .139510
/notes="assembly fragment:00225
fragment_chain:3"
misc_feature
139611. .144398
/notes="assembly fragment:00113
fragment_chain:3"
misc_feature
144499. .146650
/notes="assembly fragment:00015
fragment_chain:3"
misc_feature
146751. .149981
/notes="assembly fragment:00028
fragment_chain:4"
misc_feature
150082. .158194
/notes="assembly fragment:00329
fragment_chain:4"
misc_feature
158295. .165247
/notes="assembly fragment:00140
fragment_chain:5"
misc_feature
165348. .175979
/notes="assembly fragment:00477
fragment_chain:5"
misc_feature
176080. .179684
/notes="assembly fragment:00038"
misc_feature
179785. .185549
/notes="assembly fragment:00293.0"
misc_feature
185650. .210916
/notes="assembly fragment:01006"
```

Query Match 62.7%; Score 20.8; DB 12; Length 210916;
Best Local Similarity 52.4%; Pred. NO. 6.7e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCTGGCCNTCNCNTAYARGAYTAYGAGTNAAR 42
DB 180085 GCGTGTGCAAGCCCTCTTCTTTCAAGGCTATGATGTGAAA 180044

RESULT 74
AC096459/c
LOCUS
DEFINITION
ACCESSION
AC096459
VERSION
GTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus;
1 (bases 1 to 212324)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Dey-Carroll, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,
Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenuhewa, L., Loulseged, H., Lozado, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwokeneme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Direct Submission
Unpublished
2 (bases 1 to 212324)
Worley, K.C.
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212324)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24942666.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFWW
Center clone name: CH230-147D4
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 205118 bases at least Q40
Consensus quality: 206969 bases at least Q30
Consensus quality: 208110 bases at least Q20
Estimated insert size: 21332; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 212324: contig of 212324 bp in length.
FEATURES
source
1 .212324
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-147D4"

ORIGIN
Query Match 62.78; Score 20.8; DB 12; Length 212324;
Best Local Similarity 52.44; Pred. No. 6.7e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
QY 1 GCNCAYCNGTGGCCNTCNCNTAYARGAYTAYGARTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192354 GCACATGTGTCTCATACATTACACACAACTGAATTAAA 192313
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 75

```

AC137342/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC137342 219429 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH230-unknown, WORKING DRAFT SEQUENCE.
AC137342
AC137342.1 GI:25138433
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULFILLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus;
1 (bases 1 to 219429)

REFERENCE
AUTHORS

Muzny, D., Maric, M., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alibab, S., Amin, A., Anguiano, D.,
Aryalabechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buha, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederick, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C., Gabisi, A., Gant, R., Garcia, A., Garner, T., Garza, W.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hollins, B., Hines, S., Hladik, S., Hodgson, A., Hogues, M.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louised, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmood, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M., McNeill, T., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervile, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwankweme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parke, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Poppo, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Williams, G., Willson, R., Wleczyk, R., Woodden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D., R., Holt, R., A., Smith, H., O.,
Weinstock, G., and Gibbs, R., A.

TITLE
JOURNAL

Unpublished
2 (bases 1 to 219429)
Rat Genome Sequencing Consortium.
Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas

COMMENT

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23096320. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GFWD

Center clone name: CH230-103A13

----- Summary Statistics

Assembly program: Atlas 3.0:

Consensus quality: 192971 bases at least Q40

Consensus quality: 199252 bases at least Q30

Consensus quality: 203322 bases at least Q20

Estimated insert size: 207485; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 236414: contig of 236414 bp in length
* 236415 236514: gap of unknown length
* 236515 238064: contig of 1350 bp in length
* 238065 238164: gap of unknown length
* 238165 239372: contig of 1208 bp in length
* 239373 239472: gap of unknown length
* 239473 241557: contig of 2085 bp in length.

FEATURES Location/Qualifiers

1. 241557
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-103A13"

misc_feature 1..2015
/note="wgs contig"
misc_feature 114280..116396
/note="wgs contig"
misc_feature 166435..168513
/note="wgs contig"
misc_feature 169489..171488
/note="wgs contig"

gap 236415..236514
/estimated_length=unknown
gap 238065..238164
/estimated_length=unknown
gap 239373..239472
/estimated_length=unknown

ORIGIN

Query Match 62.7%; Score 20.8; DB 12; Length 241557;

Best Local Similarity 59.5%; Pred. No. 6.8e+02;
Matches 22; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
Qy 1 GCNCAVCNCGCCNCNCNTAARGAYATYARG 37
Db 45208 GTGCATTGTCCTTTCTTACAAAGATAGGAGG 45244

RESULT 79

AC134740

LOCUS

DEFINITION Rattus norvegicus clone CH230-81D7, WORKING DRAFT SEQUENCE, 2
unordered pieces.

AC134740.2 GI:25138728

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

1 (bases 1 to 243005)

Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Aisbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyte,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Dalgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Georgiadis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

Hollands,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorensueta,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,

Mareshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,

Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,

Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,

Reilly,B., Reilly,W., Ren,Y., Reuter,M., Richards,S., Riggs,F.,

Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,

Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,

Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,

Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,

Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,

Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,V.,

Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von

Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,

Neistack,G. and Gibbs,R.A.

Direct Submission

Unpublished

TITLE

JOURNAL

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, J., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 319548)
Worley, K.C.
Direct Submission
Submitted (16-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 319548)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23816394.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSQR
Center clone name: CH230-135F12
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 258869 bases at least Q40
Consensus quality: 263450 bases at least Q30
Consensus quality: 267288 bases at least Q20
Estimated insert size: 267899; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 13284: contig of 13284 bp in length
* 13285 13384: gap of unknown length
* 13385 33286: contig of 19902 bp in length

* 33287	33386: gap of unknown length
* 33387	36573: contig of 3187 bp in length
* 36574	36673: gap of unknown length
* 50542	50542: contig of 13869 bp in length
* 50543	50642: gap of unknown length
* 50643	278526: contig of 227884 bp in length
* 278527	278626: gap of unknown length
* 278627	289355: contig of 10729 bp in length
* 289356	289455: gap of unknown length
* 289456	299608: contig of 10153 bp in length
* 299609	299708: gap of unknown length
* 299709	310176: contig of 10458 bp in length
* 310177	310276: gap of unknown length
* 310277	311613: contig of 1337 bp in length
* 311614	311713: gap of unknown length
* 311714	313058: contig of 1345 bp in length
* 313059	313158: gap of unknown length
* 313159	315145: contig of 1987 bp in length
* 313159	315245: gap of unknown length
* 315246	317898: contig of 2653 bp in length
* 317899	317998: gap of unknown length
* 317999	319548: contig of 1550 bp in length.
FEATURES	Location/Qualifiers
source	1..319548
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-135F12"
gap	13285..13384
	/estimated_length=unknown
misc_feature	24923..26466
	/notes="wgs contig"
gap	33287..33386
	/estimated_length=unknown
misc_feature	33387..34475
	/note="wgs contig"
gap	36574..36673
	/estimated_length=unknown
gap	50543..50642
	/estimated_length=unknown
misc_feature	50643..51978
	/note="wgs contig"
misc_feature	complement(219445..220297)
	/note="clone_boundary"
	clone_end:T7
	site:
	end sequence:RAWL30TUB"
misc_feature	276518..278526
	/note="wgs_end_extension"
gap	278527..278626
	/estimated_length=unknown
misc_feature	281570..282956
	/note="wgs_end_extension"
gap	289356..289455
	/estimated_length=unknown
gap	299609..299708
	/estimated_length=unknown
misc_feature	299709..301747
	/note="wgs_end_extension"
gap	310177..310276
	/estimated_length=unknown
gap	311614..311713
	/estimated_length=unknown
gap	313059..313158
	/estimated_length=unknown
gap	315146..315245
	/estimated_length=unknown
gap	317899..317998
	/estimated_length=unknown
	ORIGIN

Query Match 62.7%; Score 20.8; DB 12; Length 319548;
 Best Local Similarity 52.4%; Pred. No. 7.3e+02;
 Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GCNCAYGNTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 191152 GCACATGTGTCCATACACTTACAACAACAAATGAATTAAA 191193

RESULT 82
 AX492783/c AX492783 349980 bp DNA linear PAT 26-SEP-2002
 LOCUS Sequence 1101 from Patent EP127152.
 DEFINITION AX492783
 ACCESSION AX492783
 VERSION AX492783.1 GI:23339466

KEYWORDS Bifidobacterium longum biovar Longum
 SOURCE Bifidobacterium longum biovar Longum
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 Bifidobacteriaceae; Bifidobacterium.

REFERENCE 1
 AUTHORS None.

TITLE Bacterial strain and genome of bifidobacterium
 JOURNAL Patent: EP 127152-A 1101 31-JUL-2002;
 SOCIETE des Produits Nestle S.A. (CH)

FEATURES
 source 1..349980
 /organism="Bifidobacterium longum biovar Longum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1679"
 /note="1106 seq in place of 1099 because seq 1 (2.256.638
 bases splitted
 in 7 more sequences.
 seq 0001: from 0.000.001 to 0.349.980 length: 349980
 seq 1100: from 0.300.001 to 0.649.980 length: 349980
 seq 1101: from 0.600.001 to 0.949.980 length: 349980
 seq 1102: from 0.900.001 to 1.249.980 length: 349980
 seq 1103: from 1.200.001 to 1.549.980 length: 349980
 seq 1104: from 1.500.001 to 1.849.980 length: 349980
 seq 1105: from 1.800.001 to 2.149.980 length: 349980
 seq 1106: from 2.100.001 to 2.256.638 length: 156638"

ORIGIN

Query Match 62.7%; Score 20.8; DB 2; Length 349980;
 Best Local Similarity 55.0%; Pred. No. 7.3e+02;
 Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GCNCAYGNTGGCCNTCCNTAYAARGAYTAYGARGTNA 40
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 129737 GCCACCACTGTCTCTCCAGTAGAGGACTGCGAGCGGA 129698

RESULT 83
 AX553950/c AX553950 349980 bp DNA linear PAT 27-NOV-2002
 LOCUS Sequence 4 from Patent WO02074798.
 DEFINITION AX553950
 ACCESSION AX553950
 VERSION AX553950.1 GI:25897903

KEYWORDS Bifidobacterium longum biovar Longum
 SOURCE Bifidobacterium longum biovar Longum
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 Bifidobacteriaceae; Bifidobacterium.

REFERENCE 1
 AUTHORS Arigoni, F., Delley, M., Mollet, B., Pridmore, R.D., Schell, M.A.,
 Pohl, T.G. and Zwielen, M.C.

TITLE The genome of a bifidobacterium
 JOURNAL Patent: WO 02074798-A 4 26-SEP-2002;
 SOCIETE DES PRODUITS NESTLE S.A. (CH)

FEATURES
 source 1..349980
 /organism="Bifidobacterium longum biovar Longum"

/mol_type="unassigned DNA"
 /db_xref="taxon:1679"
 /note="seq 1: Original length 2256646 splitted into
 seq 1: from 0.000.001 to 0.349.980
 seq 3: from 0.300.001 to 0.649.980
 seq 4: from 0.600.001 to 0.949.980
 seq 5: from 0.900.001 to 1.249.980
 seq 6: from 1.200.001 to 1.549.980
 seq 7: from 1.500.001 to 1.849.980
 seq 8: from 1.800.001 to 2.149.980
 seq 9: from 2.100.001 to 2.256.646"

ORIGIN

Query Match 62.7%; Score 20.8; DB 2; Length 349980;
 Best Local Similarity 55.0%; Pred. No. 7.3e+02;
 Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GCNCAYGNTGGCCNTCCNTAYAARGAYTAYGARGTNA 40
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 129740 GCCACCACTGTCTCTCCAGTAGAGGACTGCGAGCGGA 129701

RESULT 84
 BV639777

LOCUS S217P60461FC10.T0 Noemie Pan troglodytes troglodytes STS genomic,
 DEFINITION sequence tagged site.

ACCESSION BV639777
 VERSION BV639777.1 GI:62667747

KEYWORDS STS.
 SOURCE Pan troglodytes troglodytes

ORGANISM Pan troglodytes troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Pan.

REFERENCE 1 (bases 1 to 770)

AUTHORS Mikkelsen, T.S., Hillier, W.L., Eichler, E.E., Zody, M.C. and
 Jaffe, D.B.

TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
 Human Genome

JOURNAL Unpublished (2005)
 COMMENT

Contact: Michael C. Zody
 Broad Institute of MIT and Harvard
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172580933
 Fax: 6172580903

Email: mczody@broad.mit.edu
 Primer A: No sequence submitted
 Primer B: No sequence submitted
 STS size: 770

Protocol:
 23,021,928 chimpanzee whole genome shotgun reads were aligned to
 the Human genome NCBI
 Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,
 including Clint (Pan
 troglodytes verus), 3 other Pan troglodytes verus chimps
 (Donald, Karlén, Yvonne), 3 Pan
 troglodytes troglodytes chimps (Noemie, Masuku, Clara) and 2 chimps
 of unknown origin
 (Gon, Unknown Chimp). Common names: Pan troglodytes verus is the
 western chimp and Pan
 troglodytes troglodytes is the central chimp. To be included in
 chimpanzee SNP discovery, a
 read must be at least 500bp in length, at least 50% of its base
 calls must have Phred
 score >= 20, at least 30% of its base calls must satisfy
 SNQS(30,25) (single strand NQS, the
 base in question has Phred score >= 30, the surrounding 10 bases in
 the read have Phred
 score >= 25), and the read must have at least 200 bp SNQS(30,25)
 bases. Reads not uniquely
 placed in the genome and read pairs whose two ends were not


```
RESULT 86
LOCUS MYCPMGAB
DEFINITION Mycoplasma gallisepticum haemagglutinin homologues
            9720 bp DNA linear BCT 24-FEB-1995
            (pmgal.2-pmgal.6) genes.
ACCESSION L28424.1 GI:535687
VERSION L28424.1
KEYWORDS haemagglutinin; major surface protein.
SOURCE Mycoplasma gallisepticum
ORGANISM Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE 1 (bases 1 to 9720)
AUTHORS Markham,P.F., Grew,M.D., Sykes,J.E., Bowden,T.R., Pollocks,T.D.,
        Browning,G.F., Whithear,K.G. and Walker,I.D.
TITLE The organisation of the multigene family which encodes the major
        cell surface protein, pmGA, of Mycoplasma gallisepticum
JOURNAL FEMS Lett. 352 (3), 347-352 (1994)
PUBMED 7925999
COMMENT Original source text: Mycoplasma gallisepticum (strain S6) DNA.
FEATURES
     source              1..9720
                        /organism="Mycoplasma gallisepticum"
                        /mol_type="genomic DNA"
                        /strain="S6"
                        /db_xref="taxon:2096"
     gene                1..965
                        /gene="pmGA1.6"
     CDS                 <1..965
                        /gene="pmGA1.6"
                        /note="haemagglutinin homologue"
                        /codon_start=3
                        /transl_table=4
                        /protein_id="AAA62415.1"
                        /db_xref="GI:535688"
                        /translation="IQEFNSSTCEWIRNFKNRTKTTSWTTPAPERAAESSADS
                        NPTNQSQNAQPAESAMNETPTDGINVAKVTLTDLKFGSNTIELSVPTTDEEQT
                        SKVAPMGNVYITNDGNQRIYNEIFGNTSNDSTSVTVLLKGYSLASGWSVTGEF
                        KRLTGAQVSDSOTSPASVLYGVPIAGSGORTVNNVSNIRNPIAGQKTLAIYVNA
                        PSDQYIGSITYSNTSRSLKFGTETDDANSVITVAKATWELQTFNTSDMMVVQ
                        NNHRTUHLKQGLNKVIVSGNSDTPFFIGNLTFLLNTPSSNSSES"
                        1148..1177
     repeat_region       1148..1177
                        /note="GAA repeat"
     gene                1308..3257
                        /gene="pmGA1.2"
     CDS                 1308..3257
                        /gene="pmGA1.2"
                        /note="haemagglutinin homologue"
                        /codon_start=1
                        /transl_table=4
                        /protein_id="AAA62416.1"
                        /db_xref="GI:535689"
                        /translation="MKKKNLKFKVSLIGISFVMLAAASCTTPTPNPNPPSGGM
                        NGDTPDQGGMMNAAASQELAAARMGLTTIFDSKXNLGLYVDYKKTQNTLTAKYDA
                        AKTVLDNSSSTTNLNEAKTRLETAIRTAATSKOTPDQHAELVKVYKELKTLISNET
                        ATLAPADAOYAKMKHLSGLYDAGKAITTKLEPVEGDLTASAVMANKIVEAIK
                        DEVLNPOKENATKLADSFVQVLVKEKITGVVEAHNKAQNPANSFVGISVDITGITTF
                        QTSIPNWDYIAQRTIFTSNDEPRSISNTPADQGTWAQPLSNVSWIYSLAGTGAKTYLEF
                        TVGPSTGYLYFPYKLVNTSDQVKLEGLYKLNDAKPSAITFGSDQTMNGKTPTVNDI
                        NVAKVTLANLFGSNKIEFSPVPAEKVSPMIGNMYLSSPNWNKIYDDIFGNSVTTNQ
                        NRTIISVDALNGYSLADWSITYAEYSGAGLTLDQAKPNEKYLYIGYVGTGARNDM
                        MVPKNVQKFLANNSTNRNVYFVNAAPRSGDYIKGVFASGVGSDLKFGSTGDMSSNG
                        VTKQLFTGNLTTLTRTFDSATSTESTRVTDDPTNKKTLTLVLEGLNKIVVSGTTEING
                        APNFGYLEFILNETQPTTNVSNPS"
                        1308..1382
     sig_peptide         1308..1382
                        /gene="pmGA1.2"
     repeat_region       3521..3565
                        /note="GAA repeat"
     gene                3698..5806
                        /gene="pmGA1.3"
     CDS                 3698..5806
                        /gene="pmGA1.3"
```

```

/note="haemagglutinin homologue"
/codon_start=1
/transl_table=4
/protein_id="AAA62417.1"
/db_xref="GI:535690"
/translation="MKRKNILKFKVSLIGISFVMLAAASCTTPTPNPNPTPTPNPEP
NPGGGGMSDGNTPNGGGGMMGNDPNPNTTPEQOLAAARKLTLDLGTENTNVALY
ADYAKIQSTLSITAYMTAKTASENTSATLENLSASTTQAAIDKAANDKRVFDSVNQ
LVAYNNLTKLTKSTSLGLESENKYGKIKHLSKLFDTGSAITAKTLDPTSGERPT
LEKVEANNGIKMAISPESLKKWGNADKNEFEKNPLSKLEKASTDTANQOQPAN
WSPAAYSVDLTSNQNLPNNFAQRKVTWSENQOPGKTALVSSPVSATDVSWISLAG
EGKTYLTFEYDPDNLFLYPLKLVKAAADSSVALQYSLNKTSSKLNFKPAETVST
NTQSENEVATITSTEARSYKVLVADEAATSNMNHPTPTVSDINIAKVLGLTGF
ENTIEFSVPEKVPKPMIGNMYLTSNSQVKPIYNEIFGNTSNDSTSVTVLLTGY
SLASGWTYIGEFKQLMDTQMGASSDSQTKPSYLVGVFGPGRRLVNNKNKQSPQ
TNSQORTFLYVNAPOGDYIYSGSYLTSNRLQMFSPGDSSTSNMNVTLTQGR
NWTLLGAFNTQTNVIVNPSNLSSGMQNAHTTLRLKGLNKVIVSGVSGNDTPTIGNL
TFTLNNPSSNTSETRQST"
3698..3772
/sig_peptide
/gene="pmGA1.3"
repeat_region 5973..6014
/note="GAA repeat"
gene 6146..8254
CDS 6146..8254
/gene="pmGA1.4"
/note="haemagglutinin homologue"
/codon_start=1
/transl_table=4
/protein_id="AAA62418.1"
/db_xref="GI:535691"
/translation="MKRKNILKFKVSLIGISFVMLAAASCTTPTPNPKPNPEPKP
NPAPKDPMPNPGGMMGNDPNPNTTPEQOLAAARKLTLDLGTENTNVALY
STYADYANIQNTLITAAITAKSTSDNTSATLEQVKSATSLQTAIDTAASKTSFDEK
NPELIKAYNALKTLKKRNSLTKSDTSDNTSATLEQVKSATSLQTAIDTAASKTSFDEK
AINLSAVSQANTINISNAVKLETWKATVLTATFVKEVLVKNKLTGIDTNNQEPG
NYFGVSVDTVTCGSDNPNWSPAQKVTWTSNTDILSQPAEGENQOQAPDVSWIY
NLTMGAKYSLTFNYGPTGFYFPYKLVNSDSKVALEYKLNESAUKTIDTSPSQ
TSPVSDATRENNRSTAAPAQGSTENPAITLDDIKIAKTVLSNLKFGSNTIEFSVPT
TAKEGTSKVPAMIGNMYLTSNDRDVKNKIYDDLFGNNSVQDQNTATVVDLLKGYSL
ATNYSIVFRFMGLQENSMRTDFIYLYGIGSLDRLPNRANSKVKQFNNSPTNNN
TRFTTYIVRPMVEGVYVSGSYLFSSTQNTQRLKFLIDGNNAVSTVQRQVDMNTL
GAFDTSKTNNDQNSSVVGNIKTLRLGLKLNKIIISGGTQDQGTNAPYIGNLTFLKMT
TTSNSTNTPAEGTSTEHAK"
6146..6220
/sig_peptide
/gene="pmGA1.4"
repeat_region 8402..8431
/note="GAA repeat"
gene 8563..9720
CDS 8563..9720
/gene="pmGA1.5"
/gene="pmGA1.5"
/note="haemagglutinin homologue"
/codon_start=1
/transl_table=4
/protein_id="AAA62419.1"
/db_xref="GI:535692"
/translation="MKRKNILKFKVSLIGISFVMLAAASCTTPTPNHEPKPNPAP
KDPKPPGCGMMGNDPNPNTTPEQOLAAARKLTLDLGTENTNVALY
DYAKIKNDLTAAITTAETASQNAATLEQVKNAASTIQTATINAVNEKRVFDENNSEL
VTATYKTLTEGENTTAAFNDSANVGGIKHLLSLYNOAKTITSTLTLADAGOSNL
KONVVKINKEITDAINPTLNOQKANADMLATFTKQVLNDAOLTSGSSSETSMOTQ
PGNSFVGVSVDVTTGTSNARNPNWPAQRKWDTRNPLAQTRQSNKLTLDVSWISYLS
GMGAKYTVTFDYIGSNNAIYFPYKLVQNDNVNGLQVLLNTPKLVNFEF"
8563..8637
/sig_peptide
/gene="pmGA1.5"
ORIGIN
```

Query Match 62.0%; Score 20.6; DB 15; Length 9720;
Best local Similarity 52.6%; Pred. No. 4.7e+02;

Matches 20; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 5 AYGNTGGCCNTCCNCTAARGAVTAYGARGTNAAR 42


```

AE015450_05      500001      610000
AE015450_06      600001      710000
AE015450_07      700001      810000
AE015450_08      800001      910000
AE015450_09      900001      996422
Continuation (5 of 10) of AE015450 Mycoplasma gallisepticum s

Query Match      62.0%; Score:20.6; DB 15; Length 110000;
Best Local Similarity 52.6%; Pred. No. 7.4e+02;
Matches 20; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 5 AYGNTGGCCNTCCNTAARGAYTAYGARGNAAR 42
Db 88994 ATGCTGACAATCACCTAATAAGACAATGTAGTAAAA 89031

RESULT 90
AM050691/c
LOCUS          AM050691      112585 bp      DNA      linear      HTG 01-FEB-2006
DEFINITION    Triticum aestivum chromosome 5AL clone 606L2, *** SEQUENCING IN
PROGRESS ***, 30 unordered pieces.
ACCESSION     AM050691
VERSION       AM050691.1 GI:86439755
KEYWORDS      HTG; HTGS PHASE1.
SOURCE        Triticum aestivum (bread wheat)
ORGANISM      Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Pooleae; Triticeae; Triticum.

REFERENCE
AUTHORS       Griffiths,S., Sharp,R.R., Foote,T.N., Bertin,I., Wanous,M.,
              Reader,S. and Moore,G.
TITLE         Molecular dissection of the Ph1 locus
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 112585)
AUTHORS       Griffiths,S.
TITLE         Direct Submission
JOURNAL       Submitted (19-SEP-2005) Griffiths S., Crop Genetics, John Innes
              Centre, Norwich Research Park, Colney, Norwich, NR4 7UH, UNITED
              KINGDOM

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1354: contig of 1354 bp in length
* 1355 1454: gap of unknown length
* 1455 2800: contig of 1346 bp in length
* 2801 2900: gap of unknown length
* 2901 4245: contig of 1345 bp in length
* 4246 4345: gap of unknown length
* 4346 5714: contig of 1369 bp in length
* 5715 7159: gap of unknown length
* 7160 7259: contig of 1345 bp in length
* 7260 8576: gap of unknown length
* 8577 10015: contig of 1317 bp in length
* 10016 11636: gap of unknown length
* 11637 11727: contig of 1339 bp in length
* 11728 13952: contig of 1511 bp in length
* 13953 14052: gap of unknown length
* 14053 15566: contig of 2226 bp in length
* 15567 17035: contig of 1514 bp in length
* 17036 17135: gap of unknown length
* 17136 18227: contig of 1369 bp in length
* 18228 18327: gap of unknown length
* 18328 19508: contig of 1092 bp in length
* 19509 30449: contig of 1181 bp in length

FEATURES
source
1..112585
/organism="Triticum aestivum"
/mol_type="genomic DNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/chromosome="5AL"
/clone="606L2"
1355..1454
/estimated_length=unknown
2801..2900
/estimated_length=unknown
4246..4345
/estimated_length=unknown
5715..5814
/estimated_length=unknown
7160..7259
/estimated_length=unknown
8577..8676
/estimated_length=unknown
10016..10115
/estimated_length=unknown
11627..11726
/estimated_length=unknown
13953..14052
/estimated_length=unknown
15567..15666
/estimated_length=unknown
17036..17135
/estimated_length=unknown
18228..18327
/estimated_length=unknown
19509..19608
/estimated_length=unknown
22781..22880
/estimated_length=unknown
25388..25487
/estimated_length=unknown
30449..30548
/estimated_length=unknown

```


Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission

TITLE Submitted (05-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL**AUTHORS**

3 (bases 1 to 154471)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamaze, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE Submitted (01-JUN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL**AUTHORS**

4 (bases 1 to 154471)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamaze, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL**COMMENT**

On Dec 11, 2001 this sequence version replaced gi:14277301.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5905
 Center clone name: 183_E_24

----- Location/Qualifiers

FEATURES	source	repeat_region
1..154471	/organism="Homo sapiens"	repeat_region
/mol_type="genomic DNA"	/db_xref="taxon:9606"	repeat_region
/chromosome="15"	/map="15"	repeat_region
/clone="RP11-183E24"	/clone_lib="RPC1-11 Human Male BAC"	repeat_region
1..281	/rpt_family="MLT2B1"	repeat_region
complement(302..780)	/rpt_family="L1MEC"	repeat_region
complement(787..919)	/rpt_family="AluSp"	repeat_region
complement(924..1150)	/rpt_family="L1MEC"	repeat_region
complement(1151..1413)	/rpt_family="AluSg"	repeat_region
complement(1414..1486)	/rpt_family="L1MEC"	repeat_region
1510..1632	/rpt_family="AT rich"	repeat_region
complement(1637..1812)	/rpt_family="L1MEC"	repeat_region
2014..2155	/rpt_family="Charlie8"	repeat_region
3408..4029	/rpt_family="MLT2F"	repeat_region
complement(4506..5313)	/rpt_family="L1MA8"	repeat_region
complement(5314..5733)	/rpt_family="WMBT"	repeat_region
complement(5734..6559)	/rpt_family="L1MA8"	repeat_region
complement(6622..6823)	/rpt_family="MIR"	repeat_region
10525..10568	/rpt_family="(TG)n"	repeat_region
11002..11619	/rpt_family="L1PA14"	repeat_region
complement(12662..12984)	/rpt_family="L1MC4a"	repeat_region
13140..13335	/rpt_family="L3"	repeat_region
13354..13482	/rpt_family="(TG)n"	repeat_region
13484..13563	/rpt_family="L3"	repeat_region
complement(14210..14403)	/rpt_family="MER20"	repeat_region
complement(14450..14730)	/rpt_family="L2"	repeat_region
complement(16431..16801)	/rpt_family="L1ME3"	repeat_region
17102..17280	/rpt_family="MER5A"	repeat_region
complement(17361..17482)	/rpt_family="L2"	repeat_region
17692..17745	/rpt_family="L2"	repeat_region
complement(17786..17909)	/rpt_family="L2"	repeat_region
complement(18118..18554)	/rpt_family="MER74C"	repeat_region
19996..20304		repeat_region

```

/rpt_family="AluSx"
complement(20444..20816)
/rpt_family="MLT11"
complement(20844..20899)
/rpt_family="L3"
complement(20900..21212)
/rpt_family="AluYb8"
complement(21217..21516)
/rpt_family="AluSp"
complement(21517..21651)
/rpt_family="L3"
21651..21846
/rpt_family="MIR"
complement(22812..22893)
/rpt_family="MIR3"
24134..24320
/rpt_family="MER63A"
24956..25064
/rpt_family="MIR"
25728..26030
/rpt_family="AluJo"
26046..26636
/rpt_family="L1MC3"
complement(27077..27379)
/rpt_family="AluSx"
complement(27670..27700)
/rpt_family="L3"
28203..28304
/rpt_family="T-rich"
complement(28334..28523)
/rpt_family="HAL1"
complement(28629..28913)
/rpt_family="HAL1"
complement(29128..29461)
/rpt_family="MLT1A1"

Query Match      62.0%; Score 20.6; DB 5; Length 154471;
Best Local Similarity 57.1%; Pred. No. 7.9e+02;
Matches 20; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 4 CAYGNTGGCCNTCCNTAYARGAYTAGGAGT 38
Db 135973 CATGCTGGCCATCTCCCTTCAAGAGTCTAAAGT 135939

RESULT 93
CT025647 165165 bp DNA linear HTG 22-NOV-2005
LOCUS
DEFINITION
Danio rerio chromosome 4 clone CH211-173P13, WORKING DRAFT
SEQUENCE.
CT025647
VERSION
GI:82617448
KEYWORDS
HTG; HTGS PHASE2; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 165165)
Matthews,L.
Direct Submission
Submitted (21-NOV-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/projects/D_rerio/fags.shtml#dataeight
On Nov 22, 2005 this sequence version replaced gi:82463356.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC173P13

/rpt_family="AluSx"
complement(20444..20816)
/rpt_family="MLT11"
complement(20844..20899)
/rpt_family="L3"
complement(20900..21212)
/rpt_family="AluYb8"
complement(21217..21516)
/rpt_family="AluSp"
complement(21517..21651)
/rpt_family="L3"
21651..21846
/rpt_family="MIR"
complement(22812..22893)
/rpt_family="MIR3"
24134..24320
/rpt_family="MER63A"
24956..25064
/rpt_family="MIR"
25728..26030
/rpt_family="AluJo"
26046..26636
/rpt_family="L1MC3"
complement(27077..27379)
/rpt_family="AluSx"
complement(27670..27700)
/rpt_family="L3"
28203..28304
/rpt_family="T-rich"
complement(28334..28523)
/rpt_family="HAL1"
complement(28629..28913)
/rpt_family="HAL1"
complement(29128..29461)
/rpt_family="MLT1A1"

Query Match      62.0%; Score 20.6; DB 12; Length 165165;
Best Local Similarity 57.1%; Pred. No. 8e+02;
Matches 20; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 4 CAYGNTGGCCNTCCNTAYARGAYTAGGAGT 38
Db 99162 CATGCAATGCCATCAGCTCTACAAAGATTACCAAGT 99196

RESULT 94
AP003651/c
LOCUS
DEFINITION
Mus musculus chromosome 16 clone RP23-19A14 map 16q, *** SEQUENCING
IN PROGRESS ***, 44 unordered pieces.
AP003651
VERSION
GI:83758728
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 171116)
Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Mus musculus genomic DNA
Published Only in Database (2005)
2 (bases 1 to 171116)
Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (22-MAY-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 164875 bases at least Q40
Consensus quality: 164972 bases at least Q30
Consensus quality: 165058 bases at least Q20
Insert size: 165165; sum-of-contigs
Insert size: 176473; 13.2% error; agarose-fp
Quality coverage: 9.20x in Q20 bases; sum-of-contigs Quality
coverage: 8.72x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 165165: contig of 165165 bp in length.
Location/Qualifiers
1..165165
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="4"
/clone="CH211-173P13"
/clone_lib="CHORI-211"
1..165165
/notes="assembly_fragment:00095"
clone_end:SP6
vector_side:left
clone_end:T7
vector_side:right

misc_feature
1..165165
/notes="assembly_fragment:00095"
clone_end:SP6
vector_side:left
clone_end:T7
vector_side:right

ORIGIN
Query Match      62.0%; Score 20.6; DB 12; Length 165165;
Best Local Similarity 57.1%; Pred. No. 8e+02;
Matches 20; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 4 CAYGNTGGCCNTCCNTAYARGAYTAGGAGT 38
Db 99162 CATGCAATGCCATCAGCTCTACAAAGATTACCAAGT 99196

RESULT 94
AP003651/c
LOCUS
DEFINITION
Mus musculus chromosome 16 clone RP23-19A14 map 16q, *** SEQUENCING
IN PROGRESS ***, 44 unordered pieces.
AP003651
VERSION
GI:83758728
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 171116)
Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Mus musculus genomic DNA
Published Only in Database (2005)
2 (bases 1 to 171116)
Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (22-MAY-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

```

COMMENT

NOTE: This is a 'working draft' sequence. It currently consists of 44 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1      7288 contig of 7288 bp in length
      7389
      15472
      25490 contig of 10019 bp in length
      25531
      33522 contig of 7932 bp in length
      33623
      39902 contig of 6280 bp in length
      40003
      47009 contig of 7007 bp in length
      47110
      55409 contig of 8300 bp in length
      55510
      62347 contig of 6838 bp in length
      62448
      68266
      74517 contig of 100 bp
      80349: gap of 100 bp
      80449: gap of 100 bp
      85274: contig of 4825 bp in length
      85374: gap of 100 bp
      85375
      90641: contig of 5167 bp in length
      90642
      95715: contig of 5074 bp in length
      95716
      95815: gap of 100 bp
      95816
      99978: contig of 4163 bp in length
      99979
      100079
      105151: contig of 5073 bp in length
      105152
      105251: gap of 100 bp
      105252
      109307: contig of 4056 bp in length
      109308
      109407: gap of 100 bp
      109408
      112997: contig of 3590 bp in length
      112998
      113097: gap of 100 bp
      113098
      116822: contig of 3725 bp in length
      116823
      116922: gap of 100 bp
      116923
      121122: contig of 4200 bp in length
      121123
      121222: gap of 100 bp
      121223
      124849: contig of 3627 bp in length
      124850
      124949: gap of 100 bp
      124950
      129645: contig of 4696 bp in length
      129646
      129745: gap of 100 bp
      129746
      131801: contig of 2056 bp in length
      131802
      131901: gap of 100 bp
      131902
      142333: contig of 2332 bp in length
      142334
      143333: gap of 100 bp
      143334
      136773: contig of 2440 bp in length
      136774
      136873: gap of 100 bp
      136874
      139997: contig of 3124 bp in length
      139998
      140097: gap of 100 bp
      140098
      142968: contig of 2871 bp in length
      142969
      143068: gap of 100 bp
      143069
      145514: contig of 2446 bp in length
      145515
      145614: gap of 100 bp
      145615
      147823: contig of 2209 bp in length
      147824
      147923: gap of 100 bp
      147924
      150151: contig of 2228 bp in length
      150152
      150251: gap of 100 bp
      150252
      151924: contig of 1673 bp in length
      151925
      152024: gap of 100 bp
      152025
      153635: contig of 1611 bp in length
      153636
      153735: gap of 100 bp
      153736
      155687: contig of 1952 bp in length
      155688
      155787: gap of 100 bp
      155788
      157365: contig of 1578 bp in length
      157366
      157465: gap of 100 bp
      157466
      159194: contig of 1729 bp in length
      159195
      159294: gap of 100 bp
      159295
      160838: contig of 1544 bp in length
      160839
      160938: gap of 100 bp
      160939
      162269: contig of 1331 bp in length
      162270
      162370
      162369: gap of 100 bp
      162370
      163441: contig of 1072 bp in length
      163442
      163541: gap of 100 bp
      163542
      165037: contig of 1496 bp in length
      165038
      165137: gap of 100 bp
      165138
      166362: contig of 1225 bp in length
      166363
      166462: gap of 100 bp
      166463
      167580: contig of 1118 bp in length
      167581
      167680: gap of 100 bp
      167681
      168767: contig of 1087 bp in length
      168768
      168867: gap of 100 bp
      168868
      169871: contig of 1004 bp in length
      169872
      169971: gap of 100 bp
      169972
      171116: contig of 1145 bp in length.
```

* NOTE: This is a 'working draft' sequence. It currently consists of 44 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1      7288: contig of 7288 bp in length
      7289
      7389
      15371: contig of 7983 bp in length
      15372
      15471: gap of 100 bp
      15472
      25490: contig of 10019 bp in length
      25491
      33522: contig of 7932 bp in length
      33523
      33622: gap of 100 bp
      33623
      39902: contig of 6280 bp in length
      39903
      40002: gap of 100 bp
      40003
      47009: contig of 7007 bp in length
      47010
      47109: gap of 100 bp
      47110
      55409: contig of 8300 bp in length
      55410
```


numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 RP24-120H1 is from a Male (C57BL/6J) mouse BAC Library VECTOR:
 pTARBAC1.

```

FEATURES             Location/Qualifiers
     source           1..186376
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="16"
                     /clone="RP24-120H1"
                     /clone_lib="RPCI-24"

ORIGIN
Query Match          62.0%; Score 20.6; DB 6; Length 186376;
Best Local Similarity 52.6%; Pred.No. 8.2e+02;
Matches 20; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy  5  AYGCTGGCCTNCNCNTAYAGGAYTAYGARGTNAAR 42
      |:||||| | | | | | | | | | | | | | | | |
Db  158812 ACACCTGGCTTATCATGACAGGCTTACAAGTCANA 158775
      |:||||| | | | | | | | | | | | | | | | |

RESULT 97
LOCUS      AC123073              191693 bp    DNA    linear    ROD 08-NOV-2003
DEFINITION Mus musculus BAC clone RP23-111E4 from 6, complete sequence.
ACCESSION  AC123073
VERSION    AC123073.4  GI:28630141
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE  1  (bases 1 to 191693)
            Searangen-Shahid,S., Bielicki,L., Dignan,G. and Shahid,S.
            The sequence of Mus musculus BAC clone RP23-111E4
            Unpublished (2001)
REFERENCE  2  (bases 1 to 191693)
            Wilson,R.
            Sequencing of Mus musculus
            Unpublished (2001)
REFERENCE  3  (bases 1 to 191693)
            McPherson,J.D. and Waterston,R.H.
            Direct Submission
            Submitted (27-MAY-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  4  (bases 1 to 191693)
            McPherson,J.D. and Waterston,R.H.
            Direct Submission
            Submitted (19-FEB-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  5  (bases 1 to 191693)
            McPherson,J.D. and Waterston,R.H.
            Direct Submission
            Submitted (02-MAR-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  6  (bases 1 to 191693)
            Wilson,R.
            Direct Submission
            Submitted (08-NOV-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Mar 2, 2003 this sequence version replaced gi:28416289.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC

```

Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics

 Center project name: M_BA0111E04

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
 The RPCI-23 BAC Library has been constructed by Kazutoyo Oseigawa
 and Minako Tateno in the laboratory of Pieter de Jong
 (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
 brain genomic DNA. The clone and detailed information can be
 obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
 Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
 This sequence is the entire insert of the clone.

```

FEATURES             Location/Qualifiers
     source           1..191693
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="6"
                     /map="6"
                     /clone="RP23-111E4"
                     /clone_lib="RPCI-23"

     repeat_region    1..160
                     /rpt_family="L1"
     repeat_region    158..1380
                     /rpt_family="L1"
     repeat_region    1453..1556
                     /rpt_family="B4"
     repeat_region    2213..2410
                     /rpt_family="B4"
     repeat_region    2772..3004
                     /rpt_family="B4"
     repeat_region    3381..3809
                     /rpt_family="L1"
     repeat_region    5698..6103
                     /rpt_family="L1"
     repeat_region    6180..6395
                     /rpt_family="MER1_type"
     repeat_region    7740..7798
                     /rpt_family="B4"
     repeat_region    8320..8738
                     /rpt_family="L1"
     repeat_region    8789..8965
                     /rpt_family="B2"
     repeat_region    8970..11265
                     /rpt_family="L1"
     repeat_region    11575..12823
                     /rpt_family="L1"
     repeat_region    12826..14683
                     /rpt_family="L1"
     repeat_region    14689..14883

```

```
repeat_region /rpt_family="MaLR"
14884..16083
repeat_region /rpt_family="L1"
16092..16331
repeat_region /rpt_family="L1"
16333..16575
repeat_region /rpt_family="MaLR"
1756..19175
repeat_region /rpt_family="L1"
23186..23271
repeat_region /rpt_family="L1"
23375..23765
repeat_region /rpt_family="L1"
23766..24913
repeat_region /rpt_family="ERVK"
24926..25072
repeat_region /rpt_family="ERVK"
25089..25393
repeat_region /rpt_family="ERVK"
25824..26056
repeat_region /rpt_family="L1"
26104..27070
repeat_region /rpt_family="L1"
27063..27699
repeat_region /rpt_family="L1"
27703..27935
repeat_region /rpt_family="MaLR"
27936..28264
repeat_region /rpt_family="MaLR"
28282..28592
repeat_region /rpt_family="L1"
29879..31448
repeat_region /rpt_family="MaLR"
31460..31969
repeat_region /rpt_family="L1"
32044..32177
repeat_region /rpt_family="Alu"
32100..32194
repeat_region /rpt_family="B4"
32212..32263
repeat_region /rpt_family="L1"
33702..33767
repeat_region /rpt_family="B4"
34184..34276
repeat_region /rpt_family="Alu"
35093..35241
repeat_region /rpt_family="B4"
35658..35846
repeat_region /rpt_family="L1"
35835..36421
repeat_region /rpt_family="L1"
36750..36962
repeat_region /rpt_family="MaLR"
37098..40719
repeat_region /rpt_family="L1"
40713..42128
repeat_region /rpt_family="L1"
42729..42865
repeat_region /rpt_family="MaLR"
43147..43333
repeat_region /rpt_family="B2"
43432..43581
repeat_region /rpt_family="L1"
44196..44397
repeat_region /rpt_family="B4"
44652..44842
repeat_region /rpt_family="ERVK"
44757..44990
repeat_region /rpt_family="B4"
45301..45394
repeat_region /rpt_family="L1"
45466..45894
repeat_region /rpt_family="L1"

repeat_region 47361..47698
/rpt_family="ERVK"
54091..54444
/rpt_family="ERVK"
54988..55095
/rpt_family="MER1_type"
57030..57250
/rpt_family="L1"
57387..57797
/rpt_family="L1"
57837..58025
/rpt_family="B2"

Query Match 62.0%; Score 20.6; DB 6; Length 191693;
Best Local Similarity 64.5%; Pred. No. 8.2e+02;
Matches 20; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 8 CNTGGCNCNTCNCNTAYAAAGAYTAYGARGT 38
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172540 CTTGGCATTCCCATATAAAGAAATATAAGT 172570

RESULT 98
AC147235/c 193668 bp DNA linear ROD 27-JAN-2005
LOCUS Mus musculus BAC clone RP23-353B7 from 6, complete sequence.
DEFINITION AC147235
ACCESSION AC147235.3 GI:54292044
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 193668)
AUTHORS Zheng, X., Cotton, M., Haglund, K. and Meyer, R.
TITLE The sequence of Mus musculus BAC clone RP23-353B7
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 193668)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 193668)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 193668)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 193668)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Oct 19, 2004 this sequence version replaced gi:51315670.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
----- Center project name: M_BA0353B07
-----

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
```

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Phred information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseawa and Minako Taneno in the laboratory of Pieter de Jong (http://www.chori.org) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

```

Location/Qualifiers
1..193668
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="6"
   /map="6"
   /clone="RP23-353B7"
   /clone_lib="RPCI-23"
misc_feature
149359..149522
   /note="Sequence derived from PCR product of project DNA."

```

ORIGIN

Query Match 62.0%; Score 20.6; DB 6; Length 193668;

Best Local Similarity 64.5%; Pred. No. 8.2e+02;

Matches 20; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 8 CNTGGCCNTCCNTAAGCATATGAGT 38

Db 176270 CTTGGCATCCCCATATAAGATATAAGT 176240

RESULT 99

AC127132

DEFINITION Rattus norvegicus clone CH230-156A5, WORKING DRAFT SEQUENCE, 2 unordered pieces.

AC127132 GI:30521218

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 207858)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindarne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinec,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa M., Murphy,J., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakoelameh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,P., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartabeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,B., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 207858)

Direct Submission

Worley,K.C.

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 207858)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23664503.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: G10F

Center clone name: CH230-156A5

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 199732 bases at least Q40

Consensus quality: 201456 bases at least Q30

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:16:40 ; Search time 312.2 Seconds
(without alignments)
937.971 Million cell updates/sec

Title: US-10-600-816-33

Perfect score: 33.2

Sequence: 1 gcnaycgntggcncntcc.....ayaargaytaygargtnaar 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

N_Geneseq_8.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

14: Geneseqn2005s.*

15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33.2	100.0	42	12	ADI28530 Human GPC
2	30	90.4	497	11	ADT96418 Colon can
3	30	90.4	497	11	ADT96418 Colon can
4	30	90.4	552	11	ADT96292 Colon can
5	30	90.4	552	11	ADT96292 Colon can
6	30	90.4	603	3	AAC399079 Human cell
7	30	90.4	620	6	ABQ58527 Human col
8	30	90.4	634	6	ABQ58527 Human col
9	30	90.4	642	10	ABT22328 Breast ca
10	30	90.4	1074	12	ADO30035 Human GPC
11	30	90.4	1212	3	AAC399079 Human GPC
12	30	90.4	1212	6	ABL34861 Human cell
13	30	90.4	1212	6	ABL34861 Human cell
14	30	90.4	1228	8	ACA04775 cDNA enco
15	30	90.4	1460	14	AEA00112 Human TAT
16	30	90.4	1460	14	AEA00112 Human TAT
17	30	90.4	1619	4	AAF58615 Human REC
18	30	90.4	1718	3	AAZ90056 Hydrophob

19	30	90.4	1788	10	ADF70573	Adf70573 Orphan re
20	30	90.4	2302	6	ABT10173	Abt10173 Human bre
21	30	90.4	2302	8	ACC58386	Acc58386 Human GPC
22	30	90.4	2302	10	ADD93240	Add93240 RAIG1 cod
23	30	90.4	2302	10	ADL24773	Adl24773 Intestina
24	30	90.4	2316	10	ABT31923	Abt31923 Human bre
25	30	90.4	2446	4	AAH14688	Aah14688 Human cDN
26	30	90.4	2456	8	ABZ42832	Abz42832 Human G p
27	30	90.4	2456	10	ACC72695	Acc72695 Human can
28	30	90.4	2456	11	ADN39301	Adn39301 Cancer/an
29	30	90.4	2456	12	ADI28525	Adi28525 Human GPC
30	30	90.4	2456	12	ADI28459	Adi28459 Human GPC
31	30	90.4	2456	13	ADR48221	Adr48221 Human ret
32	30	90.4	2456	13	ACN39388	Acn39388 Tumour-as
33	30	90.4	2456	13	ADR43992	Adr43992 Human bre
34	30	90.4	2456	13	ADU06126	Adu06126 Novel bro
35	30	90.4	2456	14	AEA00088	Aea00088 Human TAT
36	30	90.4	2456	14	AEA00608	Aea00608 Human TAT
37	30	90.4	2456	14	AEQ47480	Aeq47480 Retinoleic
38	30	90.4	2593	6	ABQ54954	Abq54954 Human ova
39	30	90.4	3371	4	AAH72766	Aah72766 Human cer
40	30	90.4	4239	11	ACN89274	Acn89274 Breast ca
41	30	90.4	6730	13	ADX97494	Adx97494 Pancreat
42	27.4	82.5	680	4	AAH08354	Aah08354 Human cDN
43	21.4	64.5	167932	10	ADL13501	Adl13501 Osteoarth
44	21.2	63.9	7880	13	ADR84444	Adr84444 Aspergill
45	20.8	62.7	421	13	ADQ56529	Adq56529 Novel can
46	20.8	62.7	349980	6	ABQ81845	Abq81845 Bifidobac
47	20.6	62.0	2004	2	AAT51538	Aat51538 Mycobacte
48	20.6	62.0	2577	2	AAQ68672	Aaq68672 Partial P
49	20.4	61.4	1053	14	AED71703	Aed71703 Corynebac
50	20.4	61.4	1074	5	AAH65560	Aah65560 C glutam
51	20.4	61.4	1201	11	ADL65724	Adl65724 C. glutam
52	20.4	61.4	2780	6	ABQ76469	Abq76469 S. cerevi
53	20.4	61.4	349980	5	AAH68525	Aah68525 C glutami
54	20.2	60.8	1146	8	ACA37503	ACA37503 Prokaryot
55	20.2	60.8	110000	13	ABD32911	ABD32911 2
56	20.2	60.8	110000	14	AEB39175	Aeb39175 22
57	20.2	60.8	110000	14	AE842401	AE842401 22
58	20.2	60.8	110000	14	AE842737	AE842737 08
59	20.2	60.8	189430	14	AEB35718	Aeb35718 L. pneumo
60	20.2	60.8	263853	14	AE831971	Aeb31971 L. pneumo
61	20	60.2	669	6	ABL91353	AbL91353 Chlamydia
62	20	60.2	5161	4	AAS36060	Aas36060 Human car
63	20	60.2	5161	10	AD546754	Ad546754 Human car
64	20	60.2	5161	13	ADJ08172	Adj08172 Human car
65	20	60.2	110000	2	AAJ91990	AAJ91990 08
66	20	60.2	110000	13	ABD32968	ABD32968 6
67	19.6	59.0	356	14	ADY98830	Ady98830 T. reesei
68	19.6	59.0	691	6	ABQ55988	Abq55988 Human ova
69	19.6	59.0	1083	13	ADX63042	Adx63042 Plant ful
70	19.6	59.0	1305	13	ADX36401	Adx36401 Plant ful
71	19.4	58.4	286	10	AB290108	Ab290108 Human nuc
72	19.4	58.4	286	11	ABD26338	ABD26338 AA459692
73	19.4	58.4	354	4	AAK58278	Aak58278 Human imm
74	19.4	58.4	614	10	ACD92866	AcD92866 Human col
75	19.4	58.4	918	6	ABZ57439	Abz57439 Formamido
76	19.4	58.4	918	13	ADW28597	Adw28597 cDNA enco
77	19.4	58.4	992	13	ADT04975	Adt04975 Haemophil
78	19.4	58.4	1757	4	AAH99671	Aah99671 Human pro
79	19.4	58.4	2321	5	ABX71365	Abx71365 Human sig
80	19.4	58.4	2416	6	AB211510	Ab211510 Human pol
81	19.4	58.4	2416	12	ADM44028	Adm44028 Novel hum
82	19.4	58.4	3688	3	ACA47167	AcA47167 Arabidops
83	19.4	58.4	3738	10	ADF01135	Adf01135 Bacterial
84	19.2	57.8	418	2	AAZ11190	Aaz11190 Interleuk
85	19.2	57.8	418	6	AAD39735	Aad39735 Human int
86	19.2	57.8	548	6	AAD39736	Aad39736 Human int
87	19.2	57.8	550	2	AAZ11191	Aaz11191 Interleuk
88	19.2	57.8	1001	3	AAC57760	Aac57760 Arachdon
89	19.2	57.8	1068	12	ADO30325	Ado30325 Mouse GPC
90	19.2	57.8	1704	12	ADQ08695	Adq08695 Ciona int
91	19.2	57.8	1828	4	AAS56974	Aas56974 C. tracho

92	19.2	57.8	1828	10	ADD42768	Ad442768 Chlamydia	165	18.6	56.0	1284	8	ACA39199	Ac39199 Prokaryot
93	19.2	57.8	2090	4	ABL12999	Ab112999 Drosophil	166	18.6	56.0	1305	12	ADL03247	Ad103247 DNA encod
c 94	19.2	57.8	2496	12	ADL112908	Ad112908 Human ste	c 167	18.6	56.0	1351	14	ADZ60298	Adz60298 Murine 25
c 95	19.2	57.8	2648	4	AAS34914	Aas34914 cDNA enco	c 168	18.6	56.0	1616	13	ADX65542	Adx65542 Plant ful
c 96	19.2	57.8	2648	8	ADC46072	Adc46072 Human neo	c 169	18.6	56.0	1984	13	ADX62925	Adx62925 Plant ful
c 97	19.2	57.8	3096	8	ACA30557	Ac30557 Prokaryot	c 170	18.6	56.0	67167	14	ADY80053	Ady80053 Nucleotid
c 98	19.2	57.8	3397	13	ADV35091	Adv35091 Rat cDNA	c 171	18.6	56.0	94750	4	AAF28551	Aaf28551 Genomic f
99	19.2	57.8	3397	14	ADZ60766	Adz60766 Rat Ptpn	c 172	18.6	56.0	110000	13	ABD32721_1	Abd32721_1
100	19.2	57.8	3397	14	ADZ60766	Adz60766 Rat Ptpn	c 173	18.6	56.0	130480	4	AAF25833	Aaf25833 R. marinu
101	19.2	57.8	3477	10	ABT41712	Abt41712 Toxicity	c 174	18.6	56.0	214	5	ABA11205	Ab11205 Human ner
102	19.2	57.8	4090	4	ABL112998	Ab112998 Drosophil	c 175	18.4	55.4	320	13	ADR65193	Adr65193 Cotton cD
c 103	19.2	57.8	12222	4	AAF54045	Aaf54045 Human alp	c 176	18.4	55.4	425	5	ADL41653	Adl41653 Human ova
c 104	19.2	57.8	68193	12	ADQ97400	Adq97400 Mouse can	c 177	18.4	55.4	483	9	ACH39801	Ach39801 Human foe
c 105	19.2	57.8	82660	11	ACN45192	Acn45192 Mouse gen	c 178	18.4	55.4	483	13	ADR65192	Adr65192 Cotton cD
c 106	19.2	57.8	83493	14	ADZ13310	Adz13310 Murine ca	c 179	18.4	55.4	519	4	AAH07459	Aah07459 Human cDN
c 107	19.2	57.8	83943	13	ABD32881	Abd32881 Mouse can	c 180	18.4	55.4	588	13	ACN60374	Acn60374 Cotton gy
c 108	19.2	57.8	110000	2	AAZ01425_00	Aaz01425 Complete	c 181	18.4	55.4	588	13	ACN60374	Acn60374 Cotton gy
c 109	19.2	57.8	110000	13	ABD32968_2	Abd32968_2	c 182	18.4	55.4	761	10	ADL06700	Adl06700 Carassius
c 110	19.2	57.8	156321	14	ACG82893	Acg82893 Breast ca	c 183	18.4	55.4	887	5	AAS66149	Aas66149 DNA encod
c 111	19.2	57.8	173805	10	ADL13775	Adl13775 Osteoarth	c 184	18.4	55.4	910	12	ADL14030	Adl14030 Human cDN
c 112	19.2	57.8	322101	10	ADZ58431	Adz58431 Human PAO	c 185	18.4	55.4	1128	14	AED24357	Aed24357 Sand rat
c 113	19.2	57.8	322101	12	AD135046	Ad135046 Human pro	c 186	18.4	55.4	1600	3	AAC47340	Aac47340 Arabidops
c 114	19	57.2	184	3	AAC17478	Aac17478 Human sec	c 187	18.4	55.4	1604	3	AAC40491	Aac40491 Arabidops
c 115	19	57.2	2232	8	ACA29136	Ac29136 Prokaryot	c 188	18.4	55.4	1644	13	ADR24452	Adr24452 Breast ca
c 116	19	57.2	3165	8	ABX13154	Abx13154 Human cDN	c 189	18.4	55.4	1663	12	ADO35460	Ado35460 Novel mou
c 117	19	57.2	3915	5	AAH77995	Aah77995 Nucleotid	c 190	18.4	55.4	2325	6	ABQ69466	Abq69466 Listeria
c 118	19	57.2	4112	5	ABV24436	Abv24436 Human pro	c 191	18.4	55.4	2376	6	ABQ67859	Abq67859 Listeria
c 119	19	57.2	4136	6	ABK10102	Abk10102 Human cDN	c 192	18.4	55.4	2397	10	AAZ51688	Aaz51688 Human nuc
c 120	19	57.2	4151	5	ABV24466	Abv24466 Human pro	c 193	18.4	55.4	2485	13	ADX11038	Adx11038 Plant ful
c 121	19	57.2	109565	13	ABD33086	Abd33086 Murine ca	c 194	18.4	55.4	2520	4	AAH34235	Aah34235 Human col
c 122	19	57.2	110000	2	AAZ01425_01	Aaz01425_01	c 195	18.4	55.4	2552	6	ABX71777	Abx71777 Human dit
c 123	18.8	56.6	255	4	AAK57623	Aak57623 Human imm	c 196	18.4	55.4	2746	4	AAI59777	Aai59777 Human pol
c 124	18.8	56.6	333	4	AAK50881	Aak50881 Human cDN	c 197	18.4	55.4	2830	4	AAI57991	Aai57991 Human pol
c 125	18.8	56.6	349	4	AAI84360	Aai84360 Human pol	c 198	18.4	55.4	2843	12	ADQ84108	Adq84108 Human tnm
c 126	18.8	56.6	413	10	ADD26877	Add26877 Human adi	c 199	18.4	55.4	2845	4	AAH18012	Aah18012 Human cDN
c 127	18.8	56.6	594	6	ABNG1380	Abng1380 Human can	c 200	18.4	55.4	2963	13	ACN42579	Acn42579 Human dia
c 128	18.8	56.6	948	6	ABL35085	Ab135085 Murine cD	c 201	18.4	55.4	2967	13	ACN42578	Acn42578 Human dia
c 129	18.8	56.6	2000	11	ACL35075	Ac135075 Rice stre	c 202	18.4	55.4	3151	9	ADB17497	Adb17497 Wheat pos
c 130	18.8	56.6	2000	11	ACL37468	Ac137468 Rice stre	c 203	18.4	55.4	3151	14	ABC75721	Aec75721 Wheat Arg
c 131	18.8	56.6	2933	12	ADG36051	Adg36051 Novel mou	c 204	18.4	55.4	5868	4	AAK89398	Aak89398 Human dig
c 132	18.8	56.6	3396	2	AAO55515	Aao55515 Prpase PT	c 205	18.4	55.4	7930	4	ABL10650	Ab110650 Drosophil
c 133	18.8	56.6	5603	4	ABL12539	Ab112539 Drosophil	c 206	18.4	55.4	10316	4	AAK89399	Aak89399 Human dig
c 134	18.8	56.6	10327	4	ABL12538	Ab112538 Drosophil	c 207	18.4	55.4	18981	4	ABL04961	Ab104961 Human rep
c 135	18.8	56.6	11566	5	ABAI4380	Abai4380 Human ner	c 208	18.4	55.4	18981	4	ABL97854	Ab197854 Human tes
c 136	18.8	56.6	12000	6	ABK34015	Abk34015 Human DNA	c 209	18.4	55.4	22651	4	AAK78202	Aak78202 Human imm
c 137	18.8	56.6	12963	8	ACF62802	Acf62802 Colon can	c 210	18.4	55.4	31337	12	ADQ97031	Adq97031 Mouse can
c 138	18.8	56.6	12963	8	ACF62802	Acf62802 Colon can	c 211	18.4	55.4	81001	4	AAF30035	Aaf30035 Human apo
c 139	18.8	56.6	12963	13	ADS89640	Ad89640 Oligonucl	c 212	18.4	55.4	110000	6	ABQ69245_10	Abq69245_10
c 140	18.8	56.6	12963	13	ADS89366	Ad89366 Oligonucl	c 213	18.4	55.4	110000	6	ABQ67197_09	Abq67197_09
c 141	18.8	56.6	14362	6	ABQ67048	Abq67048 Human ang	c 214	18.4	55.4	110000	11	ACN44014_1	Acn44014_1
c 142	18.8	56.6	14708	4	AAS45514	Aas45514 Chemical	c 215	18.4	55.4	110000	14	AEA61120_3	Aea61120_3
c 143	18.8	56.6	14708	6	ABL34245	Ab134245 Human imm	c 216	18.4	55.4	110000	14	AEA61102_1	Aea61102_1
c 144	18.8	56.6	14708	6	ABL92325	Ab192325 Chemical	c 217	18.4	55.4	161671	12	ADH30232	Adh30232
c 145	18.8	56.6	14708	6	ABK28450	Abk28450 DNA trans	c 218	18.4	55.4	174318	15	AEF92655	Aef92655 Human Ras
c 146	18.8	56.6	15863	10	ADB54294	Adb54294 Pretreate	c 219	18.4	55.4	188971	12	ADL08108	Adl08108 Human gen
c 147	18.8	56.6	15863	10	ADB54166	Adb54166 Pretreate	c 220	18.4	55.4	200000	12	ADO47193	Ado47193 DNA seque
c 148	18.8	56.6	23483	13	ABD32835	Abd32835 Mouse can	c 221	18.4	55.4	202251	11	ACN44504	Acn44504 Mouse gen
c 149	18.8	56.6	106323	8	ABX14591	Abx14591 Human chl	c 222	18.4	55.4	277616	13	ABD32602	Abd32602 Human can
c 150	18.8	56.6	106323	10	ADN46845_09	Adn46845_09	c 223	18.2	54.8	171	2	AAK83259	Aak83259 Breast ca
c 151	18.8	56.6	110000	12	ADN47591_11	Adn47591_11	c 224	18.2	54.8	171	2	AAV68857	Aav68857 DNA molec
c 152	18.8	56.6	110000	12	ADN46123_09	Adn46123_09	c 225	18.2	54.8	171	3	AAK80782	Aak80782 Human bre
c 153	18.8	56.6	110000	12	ADN46123_09	Adn46123_09	c 226	18.2	54.8	171	6	AAK99628	Aak99628 Breast tu
c 154	18.8	56.6	110000	12	ADN47209_11	Adn47209_11	c 227	18.2	54.8	171	6	ABK46672	Abk46672 Human bre
c 155	18.8	56.6	110000	12	ADN46464_09	Adn46464_09	c 228	18.2	54.8	171	8	ADA11139	Ada11139 Human bre
c 156	18.8	56.6	110000	12	ADN47960_11	Adn47960_11	c 229	18.2	54.8	171	10	ADC15112	Adc15112 Human bre
c 157	18.6	56.0	23	12	ADI28532	Adi28532 Human GPC	c 230	18.2	54.8	176	4	ABA71057	Ab71057 Human foe
c 158	18.6	56.0	265	6	ABN26856	Abn26856 Human ORF	c 231	18.2	54.8	176	4	AAI51271	Aai51271 Probe #19
c 159	18.6	56.0	400	2	AAV77256	Aav77256 Staphyloc	c 232	18.2	54.8	176	4	ABA37448	Ab37448 Probe #15
c 160	18.6	56.0	438	8	ABZ53830	Abz53830 Aspergill	c 233	18.2	54.8	176	4	AAK537448	Aak537448 Human bon
c 161	18.6	56.0	725	13	ADR13190	Adr13190 Human can	c 234	18.2	54.8	176	4	AAK19339	Aak19339 Human bra
c 162	18.6	56.0	906	10	ADP14188	Adp14188 Human end	c 235	18.2	54.8	176	4	ABK44996	Abk44996 Human liv
c 163	18.6	56.0	934	13	ADK60641	Adk60641 Plant ful	c 236	18.2	54.8	176	4	ABK44996	Abk44996 Human liv
c 164	18.6	56.0	1240	13	ADX11208	Adx11208 Plant ful	c 237	18.2	54.8	176	6	ABS19574	Abs19574 Human gen

238	18.2	54.8	395	9	ACH49830	Ach49830 Human leu	311	18	54.2	878	2	AAV16672	AAV16672 Polynucle
239	18.2	54.8	410	6	AAL13639	Aal13639 Human bre	312	18	54.2	878	6	ABK15460	Abk15460 Human col
240	18.2	54.8	449	6	ABN64231	Abn64231 Human can	313	18	54.2	878	6	ABSS2542	Abss2542 Human col
241	18.2	54.8	450	11	ACN83709	Acn83709 Breast ca	c 314	18	54.2	906	13	ADX51155	Adx51155 Plant ful
242	18.2	54.8	470	4	AAL16024	Aal16024 Probe #59	315	18	54.2	1000	15	AEF24500	Aef24500 Horse mit
243	18.2	54.8	470	4	ABA58464	Aba58464 Human foe	316	18	54.2	1000	15	AEF24497	Aef24497 Horse mit
244	18.2	54.8	470	4	AAL138116	Aal138116 Probe #68	317	18	54.2	1000	15	AEF24496	Aef24496 Horse mit
245	18.2	54.8	470	4	ABA27543	Aba27543 Probe #60	318	18	54.2	1029	8	ADA69628	Ada69628 Rice gene
246	18.2	54.8	470	4	AAK32263	Aak32263 Human bon	319	18	54.2	1029	11	ACL27148	Acl27148 Rice abio
247	18.2	54.8	470	4	AAK06574	Aak06574 Human bra	320	18	54.2	1080	8	ACA51895	Aca51895 Prokaryot
248	18.2	54.8	470	4	ABES1967	Abes1967 Human liv	321	18	54.2	1080	8	ACA19692	Aca19692 Prokaryot
249	18.2	54.8	470	6	ABS07038	Abs07038 Human gen	322	18	54.2	1081	8	ACA48632	Aca48632 Prokaryot
250	18.2	54.8	623	12	ADJ43982	Adj43982 Plant cdn	323	18	54.2	1165	5	AAH76179	Aah76179 DNA encod
251	18.2	54.8	669	14	ACL65099	Acl65099 M. xanthu	324	18	54.2	1165	5	AAH76150	Aah76150 DNA encod
252	18.2	54.8	864	9	ADA29982	Ada29982 DNA encod	325	18	54.2	1191	8	ABZ56934	Abz56934 Human RGS
253	18.2	54.8	1101	9	ADA330049	Ada330049 DNA encod	326	18	54.2	1305	13	ADR60022	Adr60022 Cotton cd
254	18.2	54.8	1284	8	ACA33044	Aca33044 Prokaryot	c 327	18	54.2	1341	8	ACA30486	Acac30486 Prokaryot
255	18.2	54.8	1335	10	ADH84124	Adh84124 Enterococ	c 328	18	54.2	1467	13	ADT47812	Adt47812 Bacterial
256	18.2	54.8	1594	10	ADJ56323	Adj56323 Human cdn	329	18	54.2	1512	10	ACA56252	Aca56252 Cow signa
257	18.2	54.8	2000	10	ACC60653	Acc60653 Gene sequ	330	18	54.2	1662	6	ADT26466	Adt26466 Human kin
258	18.2	54.8	2000	10	ADK64153	Adk64153 Disease t	331	18	54.2	1632	5	AAS68142	Aas68142 DNA encod
259	18.2	54.8	2115	13	ADS46677	Ads46677 Bacterial	332	18	54.2	1662	4	AAS06702	Aas06702 Polynucle
260	18.2	54.8	4365	6	AAS94791	Aas94791 Human DNA	333	18	54.2	1662	4	AAH78797	Aah78797 Human G-p
261	18.2	54.8	4416	3	AACS59319	Aacs59319 Human sec	334	18	54.2	1662	6	ABQ86178	Abq86178 Novel hum
262	18.2	54.8	4417	3	ABL01582	Ablo1582 Human sec	335	18	54.2	1662	6	AAD26466	Aad26466 Human kin
263	18.2	54.8	29729	2	AAI131175	Aai131175 Enterococ	336	18	54.2	1662	6	AAD28071	Aad28071 Human kin
264	18.2	54.8	29729	6	ABS98970	Abs98970 Enterococ	337	18	54.2	1662	6	AAD46350	Aad46350 Human con
265	18.2	54.8	64829	14	AEA61181	Aea61181 Human FHL	338	18	54.2	1662	8	ACC44826	Acc44826 Human G p
266	18.2	54.8	80963	4	AEA61095	Aea61095 Continuation (4 of	339	18	54.2	1662	12	ADG12799	Adg12799 Human GRK
267	18.2	54.8	110000	13	ABD32629	Abd32629 Continuation (4 of	340	18	54.2	1662	13	ADG60136	Adg60136 Human G p
268	18.2	54.8	110000	14	ABE42401	Abel42401 Continuation (29 o	341	18	54.2	1683	15	ABE28185	Aee28185 Human CSI
269	18.2	54.8	112414	6	ABLA59091	Abi59091 Nucleotid	342	18	54.2	1693	15	ABE28220	Aee28220 Human CSI
270	18.2	54.8	127722	12	ADO97301	Ado97301 Human can	343	18	54.2	1701	6	ADI16737	Adi16737 Human NOV
271	18.2	54.8	152811	14	ABDL8413	Aedl8413 Fibrotic	344	18	54.2	1761	8	ABZ56943	Abz56943 Human gen
272	18.2	54.8	160671	14	ABE04683	Aee04683 Cancer-as	345	18	54.2	1821	10	ADC39221	Adc39221 Novel hum
273	18	54.2	230	11	ADT96102	Adt96102 Colon can	346	18	54.2	1950	10	ABX07405	Abx07405 S. pneumo
274	18	54.2	230	11	ADX42584	Adx42584 Human cdn	347	18	54.2	1953	8	ABZ42476	Abz42476 Streptoco
275	18	54.2	242	15	ABE28182	Aee28182 Human CSI	348	18	54.2	1953	13	ADR92277	Adr92277 Novel S.
276	18	54.2	255	2	AAT22483	Aat22483 Human gen	349	18	54.2	1953	13	ADK44045	Adk44045 Streptoco
277	18	54.2	270	5	AAS75287	Aas75287 DNA encod	350	18	54.2	1953	14	AEA56147	Aea56147 Streptoco
278	18	54.2	270	5	AAS76260	Aas76260 DNA encod	351	18	54.2	2000	8	ADA73448	Ada73448 Rice gene
279	18	54.2	270	5	AAS75573	Aas75573 DNA encod	352	18	54.2	2000	11	ACL35310	Acl35310 Rice stre
280	18	54.2	320	5	ABV47825	Abv47825 Human pro	c 353	18	54.2	2121	11	ACL27531	Acl27531 Rice abio
281	18	54.2	376	11	ADT94503	Adt94503 Colon can	354	18	54.2	2145	6	ABZ32245	Abz32245 Candida a
282	18	54.2	376	11	ADK40985	Adk40985 Human cdn	355	18	54.2	2198	8	ABS57375	Abss57375 cDNA enco
283	18	54.2	384	12	ADQ36974	Adq36974 Cell prol	356	18	54.2	2249	4	AAH78799	Aah78799 Human G-p
284	18	54.2	394	12	ADQL5644	Adql5644 Rice stre	357	18	54.2	2259	5	AAH70320	Aah70320 DNA encod
285	18	54.2	411	6	ABV97567	Abv97567 Human pan	c 358	18	54.2	2371	2	AAT99565	Aat99565 Human ago
286	18	54.2	432	6	ABQ59094	Abq59094 Human col	c 359	18	54.2	2501	14	ADV97941	Adv97941 Human tum
287	18	54.2	446	14	ADV76065	Adv76065 Human col	360	18	54.2	2742	5	AAH46102	Aah46102 Human CLC
288	18	54.2	453	9	ACH14280	Ach14280 Human adu	361	18	54.2	2742	9	ADA37185	Ada37185 Human CLC
289	18	54.2	460	14	ACL54478	Acl54478 Human col	362	18	54.2	2742	10	ADC51724	Adc51724 Human CLC
290	18	54.2	463	13	ADR60999	Adr60999 Cotton cd	363	18	54.2	2742	11	ADZ26326	Adz26326 Human rhi
291	18	54.2	481	11	ADT97273	Adt97273 Colon can	364	18	54.2	2742	12	ADI29549	Adi29549 Human CLC
292	18	54.2	481	11	ADK43755	Adk43755 Human cdn	365	18	54.2	2742	14	AEA35969	Aea35969 Novel tra
293	18	54.2	482	6	ABV97193	Abv97193 Human pan	366	18	54.2	2745	2	AAF81927	Aaf81927 Human ICA
294	18	54.2	482	11	ADT96218	Adt96218 Colon can	367	18	54.2	2763	4	AAF80429	Aaf80429 5-enolpyr
295	18	54.2	482	11	ADT94662	Adt94662 Colon can	368	18	54.2	2813	15	AAE28219	Aee28219 Human CSI
296	18	54.2	482	11	ADT96722	Adt96722 Colon can	369	18	54.2	2825	5	AAH46124	Aah46124 Human CLC
297	18	54.2	482	11	ADT96379	Adt96379 Colon can	370	18	54.2	2854	4	AAF81787	Aaf81787 Human sec
298	18	54.2	482	11	ADK43204	Adk43204 Human cdn	371	18	54.2	2854	4	AAH34879	Aah34879 Human col
299	18	54.2	482	11	ADK42700	Adk42700 Human cdn	372	18	54.2	2867	4	AAH33285	Aah33285 Human col
300	18	54.2	482	11	ADK41144	Adk41144 Human cdn	373	18	54.2	2882	12	ADJ27235	Adj27235 Human TRI
301	18	54.2	482	11	ADK42861	Adk42861 Human cdn	374	18	54.2	2983	15	ABE28221	Abe28221 Human CSI
302	18	54.2	483	11	ADT97275	Adt97275 Colon can	375	18	54.2	3007	6	AAI70801	Aai70801 Human cal
303	18	54.2	483	11	ADK43757	Adk43757 Human cdn	376	18	54.2	3007	9	ACH66411	Ach66411 Human cal
304	18	54.2	524	6	ABV89223	Abv89223 Human col	377	18	54.2	3007	14	ADW11215	Adw11215 Human cal
305	18	54.2	541	3	AAF08538	Aaf08538 Fuarium	378	18	54.2	3109	4	AAH35019	Aah35019 Human col
306	18	54.2	541	13	ADU52579	Adu52579 Fusarium	379	18	54.2	3111	2	AZ029840	Az029840 Human mem
307	18	54.2	541	14	ADZ90582	Adz90582 Fusarium	380	18	54.2	3111	2	ADAI1073	Adai1073 Human cdn
308	18	54.2	675	13	ADG60139	Adg60139 Human G p	381	18	54.2	3111	9	ADG47986	Adg47986 Human cal
309	18	54.2	876	12	ADQ58726	Adq58726 Human col	382	18	54.2	3186	8	ACC44827	Acc44827 Human G p
310	18	54.2	878	2	AAT45884	Aat45884 Human col	383	18	54.2	3186	15	ABE27527	Abe27527 Human G p

384	18	54.2	3267	4	AAS26987	Aas26987 cDNA enco	C 457	17.8	53.6	196	6	ABS17468	Abss17468 Human gen
385	18	54.2	3267	10	ADB93165	Adb93165 Human cDN	C 458	17.8	53.6	282	6	ABL70829	AbL70829 Corn tass
386	18	54.2	3311	8	AAI29502	AAi29502 C902P det	C 459	17.8	53.6	294	10	ABX85996	Abx85996 Corn ear-
387	18	54.2	3311	8	ABZ33688	Abz33688 Human col	C 460	17.8	53.6	297	8	ABZ22706	Abz22706 Human ant
388	18	54.2	3311	10	AD663139	Ad663139 Human DNA	C 461	17.8	53.6	367	4	AAI14971	Aai14971 Probe #49
389	18	54.2	3311	10	AD663138	Ad663138 Human chl	C 462	17.8	53.6	367	4	ABA56707	Abas6707 Human foe
390	18	54.2	3311	10	AD663173	Ad663173 Human chl	C 463	17.8	53.6	367	4	AAI36319	Aai36319 Probe #50
391	18	54.2	3311	10	AD663174	Ad663174 Human DNA	C 464	17.8	53.6	367	4	ABA46158	Abas46158 Human bre
392	18	54.2	3311	10	AD62775	Ad62775 Human chl	C 465	17.8	53.6	367	4	ABA26319	Abas26319 Probe #47
393	18	54.2	3311	10	AD62774	Ad62774 Human chl	C 466	17.8	53.6	367	4	AAK30362	Aak30362 Human bon
394	18	54.2	3311	10	AD627739	Ad627739 Human chl	C 467	17.8	53.6	367	4	AAK04837	Aak04837 Human bra
395	18	54.2	3311	10	AD62740	Ad62740 Human DNA	C 468	17.8	53.6	367	4	ABS30003	Abss30003 Human liv
396	18	54.2	3311	11	ADP88234	Adp88234 Colon can	C 469	17.8	53.6	367	5	AAI04739	Aai04739 Probe #47
397	18	54.2	3311	11	ADP88270	Adp88270 Human dia	C 470	17.8	53.6	367	6	ABS04957	Abss04957 Human gen
398	18	54.2	3311	11	ADP88269	Adp88269 Colon can	C 471	17.8	53.6	375	9	ACH20498	Ach20498 Human adu
399	18	54.2	3311	11	ADP88235	Adp88235 Human dia	C 472	17.8	53.6	407	8	ABX48026	Abx48026 Bovine ES
400	18	54.2	3311	12	ADO78136	Ado78136 Human CLC	C 473	17.8	53.6	429	9	ACH46718	Ach46718 Human inf
401	18	54.2	3311	14	AEA04442	Aea04442 Human cDN	C 474	17.8	53.6	433	11	ADT96318	Adt96318 Colon can
402	18	54.2	3311	14	AEF53717	Aef53717 Human cal	C 475	17.8	53.6	433	11	ADX42800	Adx42800 Human cDN
403	18	54.2	3311	15	ABE48672	Abe48672 Human CLC	C 476	17.8	53.6	459	6	ABL93699	AbL93699 Arabidops
404	18	54.2	3683	8	ABZ09910	Abz09910 Human 5'	C 477	17.8	53.6	513	4	AAH12665	Aah12665 Human cDN
405	18	54.2	4210	4	ABL07074	AbL07074 Drosophi	C 478	17.8	53.6	531	10	ADH13134	Adh13134 H influen
406	18	54.2	4569	6	ABA91304	AbA91304 Human col	C 479	17.8	53.6	531	10	ADH13136	Adh13136 H influen
407	18	54.2	7172	2	AAV52253	Aav52253 Streptoco	C 480	17.8	53.6	531	14	ASC10699	Aec10699 Haemophil
408	18	54.2	7411	10	ACA55370	AcA55370 Transform	C 481	17.8	53.6	531	14	ASC10697	Aec10697 Haemophil
409	18	54.2	10716	6	ABL33418	AbL33418 Human enz	C 482	17.8	53.6	633	11	ACL34102	AcL34102 Rice abio
410	18	54.2	17321	10	ADB81091	AdB81091 Human enz	C 483	17.8	53.6	645	12	ADI53010	Adi53010 Human chr
411	18	54.2	32892	14	AEA61109	Aea61109 Human CLC	C 484	17.8	53.6	651	14	ADW85541	Adw85541 MAP3K9 ma
412	18	54.2	36851	6	AAD28072	Aad28072 Human kin	C 485	17.8	53.6	688	4	AAH32451	Aah32451 Human olf
413	18	54.2	44990	9	ADA02711	Ada02711 Mouse Vda	C 486	17.8	53.6	717	8	ABT32194	Abt32194 Human neu
414	18	54.2	44990	10	ADB72449	AdB72449 Mouse Vda	C 487	17.8	53.6	728	4	AAH06438	Aah06438 Human cDN
415	18	54.2	44990	10	ADE95959	AdE95959 Mouse Vda	C 488	17.8	53.6	773	2	AAV34177	Aav34177 Human sec
416	18	54.2	57296	4	AAK78847	Aak78847 Human imm	C 489	17.8	53.6	773	8	ACD08048	AcD08048 cDNA enco
417	18	54.2	57296	4	AAK78170	Aak78170 Human imm	C 490	17.8	53.6	773	14	ADZ12234	Adz12234 Human sec
418	18	54.2	57296	4	AAK79364	Aak79364 Human imm	C 491	17.8	53.6	911	13	ADX59374	Adx59374 Plant ful
419	18	54.2	57296	4	AAK86799	Aak86799 Human imm	C 492	17.8	53.6	958	13	ADX59371	Adx59371 Plant ful
420	18	54.2	61103	11	ACN43886	Acn43886 Human gen	C 493	17.8	53.6	959	13	ADX59365	Adx59365 Plant ful
421	18	54.2	77932	11	ADL27149	AdL27149 Mouse gen	C 494	17.8	53.6	1040	11	ACN86579	Acn86579 Breast ca
422	18	54.2	78028	9	ADA03071	Ada03071 Mouse mCG	C 495	17.8	53.6	1137	13	ADX30414	Adx30414 Plant ful
423	18	54.2	78028	9	ADA66355	Ada66355 Mouse mCG	C 496	17.8	53.6	1166	6	ABX15078	Abx15078 cDNA enco
424	18	54.2	80276	10	ADB72809	AdB72809 Mouse mCG	C 497	17.8	53.6	1182	4	AAI05641	Aai05641 Human rep
425	18	54.2	80276	14	ADZ12738	Adz12738 Murine ca	C 498	17.8	53.6	1183	4	AAI05642	Aai05642 Human rep
426	18	54.2	98472	6	ABS54549	Abss54549 Human mul	C 499	17.8	53.6	1253	2	AAV34241	Aav34241 Human sec
427	18	54.2	98472	8	ACF62745	Acf62745 Cancer ba	C 500	17.8	53.6	1253	8	ACD08112	AcD08112 cDNA enco
428	18	54.2	98472	8	ACF62745	Acf62745 Cancer ba	C 501	17.8	53.6	1253	14	ADZ12298	AdZ12298 Human sec
429	18	54.2	98472	10	ADB87949	AdB87949 Human UGT	C 502	17.8	53.6	1352	4	ABL15899	AbL15899 Drosophi
430	18	54.2	98472	10	ADB96932	AdB96932 Human MDR	C 503	17.8	53.6	1428	3	AACT8073	AacT8073 Human can
431	18	54.2	98472	10	ADB92123	AdB92123 Human MDR	C 504	17.8	53.6	1460	4	AAZ23115	Aaz23115 DNA enco
432	18	54.2	110000	10	ABS56454_14	Continuation (15 o	C 505	17.8	53.6	1545	10	ADC24928	Adc24928 Human bre
433	18	54.2	110000	14	AEA61163_1	Continuation (2 of	C 506	17.8	53.6	1562	3	AACT7963	AacT7963 Human can
434	18	54.2	128993	8	ACF62749	Acf62749 Cancer ba	C 507	17.8	53.6	1584	9	ADB06347	AdB06347 Alloiococ
435	18	54.2	128993	8	ACF62749	Acf62749 Cancer ba	C 508	17.8	53.6	1584	9	ADB06351	AdB06351 Alloiococ
436	18	54.2	128993	10	ADB87957	AdB87957 Human UGT	C 509	17.8	53.6	1584	9	ADB06345	AdB06345 Alloiococ
437	18	54.2	128993	10	ADB96940	AdB96940 Human MDR	C 510	17.8	53.6	1584	9	ADB06349	AdB06349 Alloiococ
438	18	54.2	128993	10	ADB96940	AdB96940 Human MDR	C 511	17.8	53.6	1615	13	ADX34982	Adx34982 Plant ful
439	18	54.2	137560	12	ADP74816	Adp74816 Parapoxvi	C 512	17.8	53.6	1724	14	ADZ64417	AdZ64417 Human can
440	18	54.2	147419	6	ABK83574	Abk83574 Human cDN	C 513	17.8	53.6	1776	14	ADV43710	Adv43710 Human psy
441	18	54.2	153752	12	ADO97531	Ado97531 Human gen	C 514	17.8	53.6	1861	15	AEF02361	Aef02361 Cotton cd
442	18	54.2	160921	11	ACN44962	Acn44962 Human gen	C 515	17.8	53.6	1984	4	ABL28309	AbL28309 Drosophi
443	18	54.2	178870	10	ADL13512	AdL13512 Osteoarth	C 516	17.8	53.6	1995	6	ABN95778	Abn95778 Gene #227
444	18	54.2	192992	13	ADB32866	AdB32866 Mouse can	C 517	17.8	53.6	2000	11	ACL38065	AcL38065 Rice stre
445	18	54.2	195102	13	ADV70461	Adv70461 Chicken g	C 518	17.8	53.6	2115	8	ACA29833	AcA29833 Prokaryot
446	18	54.2	275790	14	ADZ00150	AdZ00150 Salmonell	C 519	17.8	53.6	2247	13	ADR07715	AdR07715 Full leng
447	17.8	53.6	60	6	ABN47337	Abn47337 Human spl	C 520	17.8	53.6	2318	4	AAK53300	Aak53300 Human pol
448	17.8	53.6	196	4	AAI24159	Aai24159 Probe #14	C 521	17.8	53.6	2337	14	ABE26879	Aeb26879 Pinus rad
449	17.8	53.6	196	4	ABA69281	AbA69281 Human foe	C 522	17.8	53.6	2493	12	ADO62064	Ado62064 Transcript
450	17.8	53.6	196	4	ABA54948	AbA54948 Probe #18	C 523	17.8	53.6	2523	10	ADA53709	Ada53709 Human cod
451	17.8	53.6	196	4	ABA51270	AbA51270 Human bre	C 524	17.8	53.6	2623	4	AAK52316	Aak52316 Human pol
452	17.8	53.6	196	4	ABA36204	AbA36204 Probe #14	C 525	17.8	53.6	2638	4	AAK525138	Aak525138 Drosophi
453	17.8	53.6	196	4	AAK43379	Aak43379 Human bon	C 526	17.8	53.6	2641	8	ABX34705	Abx34705 Human mdd
454	17.8	53.6	196	4	AAK17568	Aak17568 Human bra	C 527	17.8	53.6	2644	10	ADC32066	Adc32066 Human nov
455	17.8	53.6	196	4	ABSA43002	AbSA43002 Human liv	C 528	17.8	53.6	2648	10	ADB62195	AdB62195 Human cDN
456	17.8	53.6	196	5	AAI09724	Aai09724 Probe #97	C 529	17.8	53.6	2676	11	ACL28094	AcL28094 Rice abio

530	17.8	53.6	2851	14	ADZ70724	Chromosom	Adz70724	603	17.8	53.6	160820	8	ABQ76673	Androgen
531	17.8	53.6	3057	15	AEE86935	Human che	Aee86935	604	17.8	53.6	196063	13	ABD33462	Human can
532	17.8	53.6	3122	12	ADQ63658	Novel hum	Ade63658	c 605	17.8	53.6	201766	10	ADL13771	Osteoarth
533	17.8	53.6	3144	4	ABL21333	Drosophil	Abi21333	c 606	17.8	53.6	212231	11	ACN44598	Human gen
534	17.8	53.6	3234	8	ABZ76275	Notch-int	Abz76275	c 607	17.8	53.6	272022	12	ADQ97126	Human can
535	17.8	53.6	3240	4	AAH14453	Human cDN	Aah14453	c 608	17.8	53.6	304905	11	ADP75180	Human End
536	17.8	53.6	3422	4	AAH16210	Human cDN	Aah16210	c 609	17.8	53.6	349980	13	ADT05648	Haemophil
537	17.8	53.6	3724	4	ABL15898	Drosophil	Abi15898	610	17.6	53.0	40	12	ADI28527	Human GPC
538	17.8	53.6	3935	4	ABL06648	Drosophil	Abi06648	c 611	17.6	53.0	51	4	AAA89199	DNA encod
539	17.8	53.6	4083	5	ABV29639	Human pro	Abv29639	c 612	17.6	53.0	56	12	ADJ66760	Human CAI
540	17.8	53.6	4083	5	ABV23765	Human pro	Abv23765	613	17.6	53.0	186	10	ACA55771	Mouse sig
541	17.8	53.6	4094	14	ADZ64409	Human can	Adz64409	614	17.6	53.0	186	12	ADI55567	Human pol
542	17.8	53.6	4324	3	AAC60909	Human gen	Aac60909	c 615	17.6	53.0	260	10	ABX83902	Corn ear-
543	17.8	53.6	4960	4	ABL18720	Drosophil	Abi18720	c 616	17.6	53.0	273	10	ABX85253	Corn ear-
544	17.8	53.6	5279	3	RAC60913	Human squ	Aac60913	c 617	17.6	53.0	276	5	ABV07935	Human pro
545	17.8	53.6	5349	6	ABA02537	Immune re	Abi02537	c 618	17.6	53.0	296	10	ABX88439	Corn ear-
546	17.8	53.6	5412	6	ABA02536	Immune re	Abi02536	c 619	17.6	53.0	345	2	AAQ60107	Human bra
547	17.8	53.6	5641	4	ABL21332	Drosophil	Abi21332	c 620	17.6	53.0	369	5	AAAS65648	DNA encod
548	17.8	53.6	5757	8	ABX63349	Human cDN	Abe63349	621	17.6	53.0	369	5	ABV37858	Human pro
549	17.8	53.6	5833	6	ABA02535	Immune re	Abi02535	622	17.6	53.0	372	6	ABV97774	Mouse ES
550	17.8	53.6	5965	6	ABV73788	Human pap	Abv73788	c 623	17.6	53.0	379	4	AAI83788	Human pol
551	17.8	53.6	5965	9	ACD26736	Papilloma	Acd26736	624	17.6	53.0	388	4	AAKS7060	Human imm
552	17.8	53.6	6028	6	ABA02534	Immune re	Abi02534	625	17.6	53.0	393	3	AAA35154	Human ade
553	17.8	53.6	6028	6	ABV73789	Human pap	Abv73789	626	17.6	53.0	393	3	AAF21276	Human low
554	17.8	53.6	6028	9	ACD26737	Papilloma	Acd26737	627	17.6	53.0	393	10	ABZ96970	Human nuc
555	17.8	53.6	6028	10	ADF09636	Human sig	Adf09636	628	17.6	53.0	393	11	ABD20819	Human pul
556	17.8	53.6	6028	12	ADL42990	Human PRO	Adl42990	629	17.6	53.0	469	13	ACF85420	Human SIR
557	17.8	53.6	6028	12	ADQ19450	Human sof	Adq19450	630	17.6	53.0	527	13	ACN58144	Cotton gy
558	17.8	53.6	6028	13	ADP23555	PRO polyP	Adp23555	631	17.6	53.0	529	4	ABA60419	Human toe
559	17.8	53.6	6028	15	AEE88052	Human cDN	Aee88052	632	17.6	53.0	529	4	RAI40303	Probe #89
560	17.8	53.6	6099	12	ADQ23664	Human sof	Adq23664	633	17.6	53.0	529	4	AAK34584	Human bon
561	17.8	53.6	7599	13	ADR84331	Aspergill	Adr84331	634	17.6	53.0	529	4	AAK08697	Human bra
562	17.8	53.6	7760	4	ABL15894	Aspergill	Abi15894	635	17.6	53.0	529	4	ABS34357	Human liv
563	17.8	53.6	7819	10	ADF81550	Leukaemia	Adf81550	636	17.6	53.0	595	13	ACN59327	Cotton gy
564	17.8	53.6	7947	4	ABL28308	Drosophil	Abi28308	637	17.6	53.0	672	4	AAF89964	Partial n
565	17.8	53.6	8062	8	ABX34538	Human mdd	Abx34538	638	17.6	53.0	700	4	AAH92725	Human inf
566	17.8	53.6	8070	13	ACN40496	Tumour-as	Acn40496	639	17.6	53.0	700	4	AAH92726	Human inf
567	17.8	53.6	8091	2	AAV57001	Human Not	Aav57001	640	17.6	53.0	706	3	AAF13054	Aspergill
568	17.8	53.6	8091	6	ABZ34974	Human gen	Abz34974	641	17.6	53.0	706	13	ADU57095	Aspergill
569	17.8	53.6	8091	10	ADL13816	Osteoarth	Adl13816	642	17.6	53.0	706	14	ADZ5098	Aspergill
570	17.8	53.6	8091	12	ADH34620	Notch hom	Adh34620	643	17.6	53.0	710	12	ADJ66748	Muc 16 St
571	17.8	53.6	8091	12	ADJ75112	Marker ge	Adj75112	644	17.6	53.0	731	3	RAC52525	Arabidops
572	17.8	53.6	8091	12	ADN48549	Human Not	Adn48549	c 645	17.6	53.0	732	12	ADJ42949	Plant cDN
573	17.8	53.6	8091	13	ADP23495	PRO polyP	Adp23495	c 646	17.6	53.0	787	13	ADR61608	Cotton CD
574	17.8	53.6	8091	13	ADR83501	Human Not	Adr83501	c 647	17.6	53.0	788	14	AEB66160	Rice geno
575	17.8	53.6	8091	14	ADR70404	Human Not	Adr70404	c 648	17.6	53.0	810	12	ADJ42948	Plant cDN
576	17.8	53.6	8257	6	ABS70408	Human bon	Abs70408	649	17.6	53.0	894	8	AAI94653	Human neu
577	17.8	53.6	13295	4	ABL09488	Drosophil	Abi09488	650	17.6	53.0	894	8	ACA42823	Prokaryot
578	17.8	53.6	13340	2	AAZ23522	O. longis	Aax23522	c 651	17.6	53.0	1346	2	AAZ97946	Human sec
579	17.8	53.6	15743	6	ABK28395	DNA trans	Abk28395	c 652	17.6	53.0	1346	9	ADA11561	Human cDN
580	17.8	53.6	17897	10	ADB54177	Pretraeate	Adb54177	c 653	17.6	53.0	1546	13	ADX15317	Plant ful
581	17.8	53.6	17897	13	ADS89327	Oligonuc	Ads89327	654	17.6	53.0	1620	11	ACL27597	Rice abio
582	17.8	53.6	17897	15	AEP08746	Human NDR	Aef08746	655	17.6	53.0	1620	12	ADJ39701	Plant cDN
583	17.8	53.6	18753	4	ABL20682	Drosophil	Abi20682	656	17.6	53.0	1731	10	ACF71816	Photorthab
584	17.8	53.6	21354	4	AAAS46815	Tumour su	Aae46815	c 657	17.6	53.0	1760	13	ADX49274	Plant ful
585	17.8	53.6	27425	3	RAAS3978	Dmp53 tum	Aas3978	c 658	17.6	53.0	1958	6	AAD28059	Human ade
586	17.8	53.6	32167	5	ABA20621	Human ner	Abi20621	c 659	17.6	53.0	2223	10	ABZ40838	N. gonorr
587	17.8	53.6	40568	11	ACN44896	Mouse gen	Acn44896	c 660	17.6	53.0	2258	5	AAZ92412	DNA encod
588	17.8	53.6	41150	10	ADL13819	Osteoarth	Adl13819	661	17.6	53.0	2266	4	AAK71354	Human imm
589	17.8	53.6	41150	14	AED18343	Fibrotic	Aed18343	c 662	17.6	53.0	2304	13	ADX60952	Plant ful
590	17.8	53.6	44348	12	ADN48556	Human Not	Adn48556	c 663	17.6	53.0	2389	6	ABQ70840	Listeria
591	17.8	53.6	48727	4	RAK67375	Human imm	Aak67375	664	17.6	53.0	2403	2	AAZ33603	Human bre
592	17.8	53.6	52464	14	ADZ12858	Murine ca	Adz12858	c 665	17.6	53.0	2823	15	AEP08723	Human gen
593	17.8	53.6	52764	11	ACN43932	Mouse gen	Acn43932	c 666	17.6	53.0	2823	15	AEP5067	Human gen
594	17.8	53.6	70123	11	ACN44254	Human gen	Acn44254	667	17.6	53.0	3108	11	ACN44331	Human mRN
595	17.8	53.6	96594	9	ADA02756	Human MOR	Ada02756	668	17.6	53.0	3126	8	ACA26007	Prokaryot
596	17.8	53.6	96594	10	ADB72494	Human MOR	Adb72494	669	17.6	53.0	3169	3	AAA35156	Human ade
597	17.8	53.6	96594	10	ADC85236	Human MOR	Adc85236	670	17.6	53.0	3169	3	AAF21278	Human low
598	17.8	53.6	96594	12	ADM74351	Human car	Adm74351	671	17.6	53.0	3169	10	ABZ96972	Human nuc
599	17.8	53.6	110000	2	AAT42063_01	Continuation (2 of	Adt42063_01	672	17.6	53.0	3169	11	ABD20821	Human pul
600	17.8	53.6	110000	9	ADB12064_00	Allioecoc	Adb12064_00	673	17.6	53.0	3446	12	ADO60053	CRH signa
601	17.8	53.6	110000	14	AEA61160_1	Continuation (2 of	Adt61160_1	674	17.6	53.0	3561	2	AAQ55970	PTase PT
602	17.8	53.6	110079	13	ADV35020	Murine cD	Adv35020	675	17.6	53.0	3566	6	AB199237	Mouse isc

676	17.6	53.0	3748	6	ABK34630	Abk34630 Human cDN	749	17.4	52.4	364	4	AA118044	Aa118044 Human bre
677	17.6	53.0	3804	4	AAH16044	Aah16044 Human cDN	750	17.4	52.4	415	11	ACN92660	Acn92660 Breast ca
678	17.6	53.0	4020	6	ABQ54212	Abq54212 Human ova	751	17.4	52.4	420	6	ABL78873	Ab178873 Human ova
679	17.6	53.0	4047	13	ADP24470	Adp24470 PRO polyP	752	17.4	52.4	457	13	ACF84030	Acf84030 Human SIR
680	17.6	53.0	4047	15	AE883887	Aee83887 Human cDN	753	17.4	52.4	489	5	AAH81814	Aah81814 Human dif
681	17.6	53.0	4091	4	AA158457	Rai58457 Human pol	754	17.4	52.4	489	5	AAH81638	Aah81638 Human dif
682	17.6	53.0	4091	5	ADQ98669	Adq98669 DNA encod	755	17.4	52.4	511	3	AAA43724	Aaa43724 Mouse sec
683	17.6	53.0	4091	9	ADB48429	Adb48429 Novel hum	756	17.4	52.4	516	15	AEF78593	Aef78593 Mouse gen
684	17.6	53.0	4093	4	AA160243	Rai60243 Human pol	757	17.4	52.4	518	9	ACH28318	Ach28318 Human adu
685	17.6	53.0	4115	4	AA502404	Ras02404 Human sec	758	17.4	52.4	526	12	ADN13627	Adn13627 Human pro
686	17.6	53.0	4115	8	ABZ73366	Abz73366 Secreted	759	17.4	52.4	533	6	ABQ36864	Abq36864 Oligonucl
687	17.6	53.0	4115	10	ADC20115	Adc20115 Human sec	760	17.4	52.4	532	6	ABQ36865	Abq36865 Oligonucl
688	17.6	53.0	4115	10	ABT16788	Abt16788 Human sec	761	17.4	52.4	546	11	ACN86819	Acn86819 Breast ca
689	17.6	53.0	4115	10	ABZ66974	Abz66974 Human sec	762	17.4	52.4	555	5	ABV53410	Abv53410 Human pro
690	17.6	53.0	4128	10	ADF81727	Adf81727 Leukaemia	763	17.4	52.4	586	4	AA107903	Aal07903 Human bre
691	17.6	53.0	4290	5	ABL52917	Ab152917 2-keto-D-	764	17.4	52.4	618	6	ABK33035	Abk33035 DNA encod
692	17.6	53.0	4514	10	ADB56172	Adb56172 Toxicity-	765	17.4	52.4	618	12	ADQ92260	Adq92260 Human aut
693	17.6	53.0	4514	10	ADB53674	Adb53674 Primary r	766	17.4	52.4	621	3	AAAC2900	Aac2900 Arabidops
694	17.6	53.0	4514	14	AEC11769	Aeci1769 Rat surro	767	17.4	52.4	652	5	AA505408	Aas05408 Mammalian
695	17.6	53.0	4668	3	AACT77483	Rac77483 Human ORF	768	17.4	52.4	828	6	ABK74586	Abk74586 Bacillus
696	17.6	53.0	4675	14	ADY18887	Ady18887 DNA encod	769	17.4	52.4	845	5	AA587938	Aas87938 DNA encod
697	17.6	53.0	5142	10	ADF90720	Adf90720 Human hep	770	17.4	52.4	937	11	ACH98632	Ach98632 Klebsiell
698	17.6	53.0	5142	12	ADN05411	Adn05411 Antipsori	771	17.4	52.4	955	6	ABN98556	Abn98556 Arabidops
699	17.6	53.0	5142	13	ADP55825	Adp55825 Human PRO	772	17.4	52.4	955	10	ADD35246	Add35246 Mouse mit
700	17.6	53.0	7095	14	ADP58452	Adp58452 Leishmani	773	17.4	52.4	1008	8	ACA30213	Aca30213 Prokaryot
701	17.6	53.0	7565	12	ADK70703	Adk70703 Murine CR	774	17.4	52.4	1143	3	AAA59157	Aas59157 cDNA enco
702	17.6	53.0	7657	4	AA545478	Aas45478 Chemicall	775	17.4	52.4	1143	3	AAA58985	Aas58985 cDNA enco
703	17.6	53.0	7657	6	ABL34023	Ab134023 Human imm	776	17.4	52.4	1143	10	ADB31999	Adb31999 Mouse cDN
704	17.6	53.0	14328	4	AAK71326	Rak71326 Human imm	777	17.4	52.4	1143	12	ADR13801	Adr13801 Murine in
705	17.6	53.0	14328	4	AAK71325	Rak71325 Human imm	778	17.4	52.4	1215	13	ADS46182	Ads46182 Bacterial
706	17.6	53.0	14328	4	AAK71335	Rak71335 Human imm	779	17.4	52.4	1221	3	AAZ52198	Aaz52198 Murine tr
707	17.6	53.0	14328	4	AAK71336	Rak71336 Human imm	780	17.4	52.4	1221	10	AAD54508	Aad54508 Mouse int
708	17.6	53.0	15000	3	ABN97977	Abn97977 Human ret	781	17.4	52.4	1221	12	ADQ92005	Adq92005 Mouse tra
709	17.6	53.0	16878	3	AAA81515	Raa81515 N. mening	782	17.4	52.4	1228	13	ADT17280	Adt17280 Plant cDN
710	17.6	53.0	21234	6	AAAD32039	Rad32039 Human kin	783	17.4	52.4	1258	4	AAD10108	Aad10108 Mouse zCY
711	17.6	53.0	23615	14	AEC09537	Rec09537 Human Bli	784	17.4	52.4	1258	8	ACF75039	Acf75039 Staphyloc
712	17.6	53.0	25837	4	AAK59552	Rak59552 Human imm	785	17.4	52.4	1284	6	ABN90749	Abn90749 Staphyloc
713	17.6	53.0	28983	14	AE9B96521	Aeb96521 Human ADA	786	17.4	52.4	1284	13	ADS02816	Adso2816 Scaphyloc
714	17.6	53.0	36741	4	AA503388	Ras03388 Human ade	787	17.4	52.4	1318	3	AAAC48785	Aac48785 Arabidops
715	17.6	53.0	36741	6	ABQ88153	Abq88153 Human ost	788	17.4	52.4	1319	3	AAAC38564	Aac38564 Arabidops
716	17.6	53.0	36741	6	ABS69890	Abs69890 Human ade	789	17.4	52.4	1356	10	ADB69774	Adb69774 C. neofor
717	17.6	53.0	36741	6	ABS52806	Abs52806 DNA encod	790	17.4	52.4	1425	3	AAAC38512	Aac38512 Arabidops
718	17.6	53.0	36741	15	AEF10051	Aef10051 Human ade	791	17.4	52.4	1427	3	AAAC48792	Aac48792 Arabidops
719	17.6	53.0	38920	12	ADQ97252	Adq97252 Mouse can	792	17.4	52.4	1469	13	ADT14803	Adt14803 Plant cDN
720	17.6	53.0	39325	4	AAK81660	Rak81660 Human imm	793	17.4	52.4	1505	10	ADB69413	Adb69413 C. neofor
721	17.6	53.0	40046	6	ABL68361	Ab168361 Kidney ca	794	17.4	52.4	1545	2	AAT90510	Aat90510 DNA encod
722	17.6	53.0	40046	6	ABL68822	Ab168822 Kidney ca	795	17.4	52.4	1559	3	AAAC48790	Aac48790 Arabidops
723	17.6	53.0	40046	6	ABL68140	Ab168140 Kidney ca	796	17.4	52.4	1562	3	AAAC38506	Aac38506 Arabidops
724	17.6	53.0	41685	11	ACN44330	Acn44330 Human gen	797	17.4	52.4	1592	6	ABL61903	Ab161903 Colon ade
725	17.6	53.0	42123	11	ACN44906	Acn44906 Human gen	798	17.4	52.4	1592	6	ABL64093	Ab164093 Breast ca
726	17.6	53.0	53131	14	AEA02956	Aea02956 Human pai	799	17.4	52.4	1592	6	ABT11013	Abt11013 Human bre
727	17.6	53.0	54550	9	ACF05021	Acf05021 Human cat	800	17.4	52.4	1659	13	ADT87784	Adt87784 Plant cDN
728	17.6	53.0	58922	13	ABD33407	Abd33407 Human can	801	17.4	52.4	1674	10	ACG61917	Acc61917 Gene sequ
729	17.6	53.0	62154	12	ADO97611	Adg97611 Human can	802	17.4	52.4	1674	10	ADK64875	Adk64875 Disease t
730	17.6	53.0	69652	13	ABD33315	Abd33315 Human can	803	17.4	52.4	1674	12	ADJ64177	Adj64177 S. cerevi
731	17.6	53.0	79860	11	ACN44122	Acn44122 Human gen	804	17.4	52.4	1692	5	AA580345	Aas80345 DNA encod
732	17.6	53.0	105184	6	ABK24122	Abk24122 Bacterial	805	17.4	52.4	1692	5	AA580333	Aas60333 DNA encod
733	17.6	53.0	110000	3	AAA81490	Continuation (10 o	806	17.4	52.4	1718	8	ACA34394	ACA34394 Prokaryot
734	17.6	53.0	110000	3	ACF67367	Continuation (51 o	807	17.4	52.4	1728	10	ADN95347	Adn95347 Human BEC
735	17.6	53.0	110000	10	ACP65387	Continuation (2 of	808	17.4	52.4	1897	11	ADN95347	Adn95347 Human BEC
736	17.6	53.0	110000	11	ADN27081	Continuation (2 of	809	17.4	52.4	1917	13	ADS78310	Adt78310 Aminotran
737	17.6	53.0	113575	11	ACN44348	Acn44348 Mouse gen	810	17.4	52.4	1921	13	AA126706	AA126706 Human bre
738	17.6	53.0	147708	6	ABQ88154	Abq88154 Human ost	811	17.4	52.4	1928	11	ACN88684	Acn88684 Breast ca
739	17.6	53.0	170001	14	AED18493	Aed18493 Fibrotic	812	17.4	52.4	1981	5	AA583523	Aas83523 DNA encod
740	17.6	53.0	176771	12	ADQ97167	Adg97167 Human can	813	17.4	52.4	2000	11	ACL35768	ACL35768 Rice stre
741	17.6	53.0	193672	10	ADL13570	Adl13570 Osteoarth	814	17.4	52.4	2079	2	AAT90513	Aat90513 DNA encod
742	17.6	53.0	209484	11	ACN44126	Acn44126 Human gen	815	17.4	52.4	2143	5	AA584876	Aas4876 DNA encod
743	17.6	53.0	349980	3	AAF21609	Aaf21609 Neisseria	816	17.4	52.4	2212	8	AA531176	Aas31176 Human kin
744	17.6	53.0	349980	3	AAF21608	Aaf21608 Neisseria	817	17.4	52.4	2247	5	AA569327	Aas69327 DNA encod
745	17.4	52.4	295	3	AAA87092	Raa87092 Rat hepat	818	17.4	52.4	2569	10	ADD44980	Add44980 Rat gene
746	17.4	52.4	315	10	ACD95013	Acd95013 Human col	819	17.4	52.4	2569	10	ADD44980	Add44980 Rat gene
747	17.4	52.4	350	4	AA125649	Ral25649 Human bre	820	17.4	52.4	2697	4	ABL07429	Ab107429 Drosophil
748	17.4	52.4	350	6	ABV96297	Abv96297 Human pan	821	17.4	52.4	2876	11	ADM02841	Adm02841 Human cDN

C 822	17.4	52.4	2876	14	AEC85771	Aec85771 Human CDN	C 895	17.4	52.4	165799	14	ADZ13004	Adz13004 Murine ca
C 823	17.4	52.4	2889	9	ABL16704	Ab116704 Drosophila	C 896	17.4	52.4	171162	14	AED89398	Aed89398 Human bre
C 824	17.4	52.4	2932	4	ADN97572	Adn97572 Rat presy	C 897	17.4	52.4	194266	14	AED76144	Aed76144 Mouse CA
C 825	17.4	52.4	3112	10	ADB58545	Adb58545 Toxicity-	C 898	17.4	52.4	235033	2	AAV57926	Aav57926 Hereditar
C 826	17.4	52.4	3305	10	ADD29569	Add29569 Human tum	C 899	17.4	52.4	236303	4	AAV11614	Aav11614 Human gen
C 827	17.4	52.4	3305	10	ADD29624	Add29624 Human tum	C 900	17.4	52.4	237326	2	AAV57903	Aav57903 Hereditar
C 828	17.4	52.4	3306	3	AAA87701	Aaa87701 Human sec	C 901	17.4	52.4	246386	14	AAE05008	Aae05008 Cancer-as
C 829	17.4	52.4	3306	8	ADA40549	Ada40549 Human sec	C 902	17.4	52.4	289190	13	ABD33143	Abd33143 Murine ca
C 830	17.4	52.4	3306	8	ACC50852	Acc50852 Human sec	C 903	17.4	52.4	349980	13	ADT05649	Adt05649 Haemophil
C 831	17.4	52.4	3306	8	ABZ71477	Abz71477 Secreted	C 904	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 832	17.4	52.4	3306	9	ADB91444	Adb91444 Human sec	C 905	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 833	17.4	52.4	3306	10	ADC73984	Adc73984 Human sec	C 906	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 834	17.4	52.4	3306	10	ADA56695	Ada56695 Gene enco	C 907	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 835	17.4	52.4	3308	3	AAA87681	Aaa87681 Human sec	C 908	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 836	17.4	52.4	3308	8	ADA40228	Ada40228 Human sec	C 909	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 837	17.4	52.4	3308	8	ACC50671	Acc50671 Human sec	C 910	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 838	17.4	52.4	3308	8	ABZ71365	Abz71365 Secreted	C 911	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 839	17.4	52.4	3308	9	ADB91306	Adb91306 Human sec	C 912	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 840	17.4	52.4	3308	10	ADC73754	Adc73754 Human sec	C 913	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 841	17.4	52.4	3308	10	ADA56388	Ada56388 Gene enco	C 914	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 842	17.4	52.4	3447	13	ADX33598	Adx33598 Plant ful	C 915	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 843	17.4	52.4	3505	10	ADB69052	Adb69052 C. neofo	C 916	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 844	17.4	52.4	3516	2	AAV74546	Aav74546 Staphyloc	C 917	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 845	17.4	52.4	3524	4	ABL07298	Ab107298 Drosophila	C 918	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 846	17.4	52.4	3899	6	ABI99205	Abi99205 Mouse isc	C 919	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 847	17.4	52.4	5048	5	AAST1638	Aas71638 DNA enco	C 920	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 848	17.4	52.4	5174	13	ADT05447	Adt05447 Haemophil	C 921	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 849	17.4	52.4	5443	14	ADM94163	Adm94163 Staphyloc	C 922	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 850	17.4	52.4	5565	14	ADM41323	Adm41323 Human myo	C 923	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 851	17.4	52.4	5885	15	AEF18698	Aef18698 Human KIA	C 924	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 852	17.4	52.4	6000	13	ADY81055	Ady81055 Rice anti	C 925	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 853	17.4	52.4	6126	10	ADC13546	Adc13546 Human NOV	C 926	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 854	17.4	52.4	6151	6	ABZ11646	Abz11646 Human pol	C 927	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 855	17.4	52.4	6151	12	ADM44164	Adm44164 Novel hum	C 928	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 856	17.4	52.4	6693	4	ABL07428	Ab107428 Drosophila	C 929	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 857	17.4	52.4	6881	6	ABL33381	Ab133381 Human imm	C 930	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 858	17.4	52.4	6881	6	ABK28250	Abk28250 DNA trans	C 931	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 859	17.4	52.4	8280	3	AAA95949	Aaa95949 Human KIK	C 932	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 860	17.4	52.4	8280	12	ADN10931	Adn10931 Human kal	C 933	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 861	17.4	52.4	10848	12	ADL81731	Adl81731 P. aerugi	C 934	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 862	17.4	52.4	12551	4	ABL19156	Ab119156 Drosophila	C 935	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 863	17.4	52.4	14766	4	ABL05848	Ab105848 Drosophila	C 936	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 864	17.4	52.4	19338	14	AEA61189	Aea61189 Human DNA	C 937	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 865	17.4	52.4	23377	11	ACNA44808	Acna44808 Mouse gen	C 938	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 866	17.4	52.4	25599	4	ABL16668	Ab116668 Drosophila	C 939	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 867	17.4	52.4	35510	13	ADM39407	Adm39407 Herpes si	C 940	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 868	17.4	52.4	44400	4	AAST14348	Aas14348 DNA enco	C 941	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 869	17.4	52.4	46275	6	ABT10145	Abt10145 Human bre	C 942	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 870	17.4	52.4	46275	10	ADL13621	Adl13621 Osteoarth	C 943	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 871	17.4	52.4	49826	11	ACNA44340	Acna44340 Mouse gen	C 944	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 872	17.4	52.4	60327	9	ADA02681	Ada02681 Mouse dus	C 945	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 873	17.4	52.4	60327	10	ADB72419	Adb72419 Mouse dus	C 946	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 874	17.4	52.4	60327	10	ADE95929	Ade95929 Mouse dus	C 947	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 875	17.4	52.4	81139	12	ADQ59189	Adq59189 MSI-H car	C 948	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 876	17.4	52.4	83698	6	ABN85767	Abn85767 Arabidops	C 949	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 877	17.4	52.4	89873	13	ABD32846	Abd32846 Mouse can	C 950	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 878	17.4	52.4	91080	6	ABX08336_16	Continuation (17 o	C 951	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 879	17.4	52.4	91134	12	ADO50281_16	Continuation (17 o	C 952	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 880	17.4	52.4	91134	12	ADN97989_16	Continuation (17 o	C 953	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 881	17.4	52.4	91139	12	ADJ25985_16	Continuation (17 o	C 954	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 882	17.4	52.4	91140	14	ABE85185_16	Continuation (17 o	C 955	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 883	17.4	52.4	110000	2	AAAT42063_12	Continuation (13 o	C 956	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 884	17.4	52.4	110000	6	ABQ65245_05	Continuation (6 of	C 957	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 885	17.4	52.4	110000	6	ABQ65245_06	Continuation (7 of	C 958	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 886	17.4	52.4	110000	6	ABQ67197_05	Continuation (6 of	C 959	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 887	17.4	52.4	127369	11	ACNA44006	Acna44006 Human gen	C 960	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 888	17.4	52.4	127943	12	ADQ97651	Adq97651 Human ca	C 961	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 889	17.4	52.4	127943	12	ADQ97651	Adq97651 Human gen	C 962	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 890	17.4	52.4	135928	15	AEF75216	Aef75216 Human pol	C 963	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 891	17.4	52.4	150481	14	AED89397	Aed89397 Human bre	C 964	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 892	17.4	52.4	152261	14	AED07415	Aed07415 DNA of HS	C 965	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 893	17.4	52.4	152261	14	AED00550	Aed00550 Herpes so	C 966	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence
C 894	17.4	52.4	165221	11	ACNA44524	Acna44524 Mouse gen	C 967	17.4	52.4	349980	13	AAV30053	Aav30053 Sequence

968 17.2 51.8 883 10 ADC86650 Human GPC
 c 969 17.2 51.8 903 5 AAS81735
 Ado28921 Mouse nov
 970 17.2 51.8 918 12 AD028921
 Adt15659 Plant cdn
 971 17.2 51.8 947 13 ADT15659
 Acf72582 Staphyloc
 972 17.2 51.8 972 8 ACF72582
 Adx50470 Plant ful
 973 17.2 51.8 1020 13 ADX50470
 c 974 17.2 51.8 1057 14 AEB66373
 Aeb66373 Rice geno
 975 17.2 51.8 1072 11 ACN81830
 c 976 17.2 51.8 1182 10 ACN81830
 Adc71369 Human col
 977 17.2 51.8 1282 10 ADE79051
 c 978 17.2 51.8 1296 10 ADE79038
 c 979 17.2 51.8 1299 10 ADE09872
 c 980 17.2 51.8 1322 6 ABK34963
 c 981 17.2 51.8 1330 2 AAX06100
 c 982 17.2 51.8 1335 2 AAQ81700
 c 983 17.2 51.8 1337 10 ADE79052
 c 984 17.2 51.8 1344 13 ADX52010
 c 985 17.2 51.8 1380 4 AAK16799
 c 986 17.2 51.8 1380 4 ABS42183
 c 987 17.2 51.8 1380 5 AAI08958
 c 988 17.2 51.8 1407 10 ADE79039
 c 989 17.2 51.8 1409 10 ADE79066
 c 990 17.2 51.8 1424 6 ABQ34623
 c 991 17.2 51.8 1424 6 ABQ34622
 c 992 17.2 51.8 1448 10 ADE79040
 c 993 17.2 51.8 1464 6 ABZ15412
 c 994 17.2 51.8 1475 10 ADF81935
 c 995 17.2 51.8 1503 4 AAF32484
 c 996 17.2 51.8 1549 4 ABL16991
 c 997 17.2 51.8 1574 6 RAD30575
 c 998 17.2 51.8 1591 6 ACA47751
 c 999 17.2 51.8 1616 14 ADZ81525
 1000 17.2 51.8 1623 2 AAZ96344

ALIGNMENTS

RESULT 1
 ADI28530
 ID ADI28530 standard; DNA; 42 BP.
 AC ADI28530;
 XX
 XX 22-APR-2004 (first entry)
 DT Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA.
 DE
 DE Human GPCR retinoic acid induced 3 (RAI-3) peptide DNA.
 XX
 XX Retinoic acid induced 3; RAI-3; human; G-protein coupled receptor; GPCR;
 KW antiinflammatory; immunosuppressive; cytostatic; cardiant; antiallergic;
 KW broncholytic; gene therapy; ss.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX WO2004001060-A2.
 PN
 XX
 XX 31-DEC-2003.
 XX
 XX 20-JUN-2003; 2003WO-US019255.
 XX
 XX 20-JUN-2002; 2002US-0390850P.
 PR 29-AUG-2002; 2002US-0407006P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 XX Whitney GS, Opitck G, Garulacan L, Ramanathan CS, McKinnon M;
 PI Bennett KL, Barber LE, Cacace A, Tsuchihashi Z;
 XX
 XX WPI; 2004-090973/09.
 XX
 XX New nucleic acid molecule encoding a human G-protein coupled receptor.
 PT RAI3, useful for preventing, treating or ameliorating chronic obstructive

PT pulmonary disease (COPD), COPD-like disorder, or the underlying symptoms
 of COPD.

XX Disclosure; SEQ ID NO 33; 301pp; English.

XX The present sequence is that of DNA encoding a peptide corresponding to
 CC amino acids 340-353 of a human G-protein coupled receptor, denoted
 CC retinoic acid induced 3 (RAI-3) protein ADI28460. The peptide was
 CC identified in studies of RAI-3 phosphorylation in response to cigarette
 CC smoke. Primers based on this sequence can be used in a PCR to identify
 CC individuals at risk for developing chronic obstructive pulmonary disease
 CC (COPD). RAI-3 is tyrosine phosphorylated and/or is associated/complexed
 CC with tyrosine phosphorylated proteins only in cells that have been
 CC exposed to cigarette smoke. Since RAI-3 is primarily expressed in lung
 CC tissue, and since cigarette smoke is a major causative factor of COPD,
 CC RAI-3 provides a novel cellular target for identifying modulators, e.g.
 CC agonists or antagonists, useful for the treatment and/or prevention of
 CC COPD and related disorders such as emphysema and chronic bronchitis. RAI-
 CC 3 modulators, e.g. agonists and antagonists, especially antisense
 CC compounds, can be used to treat COPD and other disorders an diseases
 CC associated with regulation of NF-kB and/or its associated or interacting
 CC signaling molecules. Single nucleotide polymorphisms (SNPs) detected in
 CC the RAI-3 gene are useful for determining COPD association in
 CC individuals. RAI-3 nucleic acid molecules and polypeptides are useful for
 CC preventing, treating or ameliorating disorders related to aberrant GPCR
 CC signaling or cell cycle regulation, pulmonary disorders, inflammatory
 CC lung disorders, COPD, the underlying symptoms of COPD, COPD-related
 CC disorders or conditions, autoimmune disorders, disorders related to
 CC hyperimmune activity, inflammatory conditions, disorders related to
 CC aberrant acute phase responses, hypercongenital conditions, birth
 CC defects, necrotic lesions, wounds, organ transplant rejection, renal
 CC diseases, ischaemia-reperfusion injury, heart disorders, disorders
 CC related to aberrant signal transduction, proliferation disorders,
 CC cancers, HIV infection, asthma, cystic fibrosis, pulmonary fibrosis,
 CC ulcerative colitis, cerebral infarct, myocardial infarct, diabetic
 CC nephropathy, allergic rhinitis, Crohn's disease, atherosclerosis,
 CC rheumatoid arthritis, inflammatory/autoimmune disorders, glioblastoma,
 CC pulmonary small cell undifferentiated carcinoma, carcinoma of the breast,
 CC colon, lung, ovary, pancreas, prostate, non-Hodgkin's lymphoma, disorders
 CC associated with aberrant cell adhesion, I-CAM function and/or regulation,
 CC E-selectin function and/or regulation, or aberrant NF-kB function and/or
 CC regulation (all claimed).

XX Sequence 42 BP; 9 A; 8 C; 7 G; 5 T; 0 U; 13 Other;

Query Match 100.0%; Score 33.2; DB 12; Length 42;

Best Local Similarity 100.0%; Pred. No. 4.2e-05;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCNTGCCNTCCNTAYARGAYTAYGARGCTNAAR 42

Dd 1 GCNCAYCNCNTGCCNTCCNTCCNTAYARGAYTAYGARGCTNAAR 42

RESULT 2

ADT96418

ID ADT96418 standard; cDNA; 497 BP.

XX AC ADT96418;

XX 16-DEC-2004 (first entry)

DT Colon cancer associated human cDNA sequence #1925.

DE
 DE Colon cancer; T cell; tumour protein; C6345; C6355; C637S; C640S; C636S;
 KW humoral immune response; cellular immune response; cytostatic;
 KW immunostimulant; human; ss.

XX Homo sapiens.

XX US2003087818-A1.

XX 08-MAY-2003.


```
PR 16-AUG-2001; 2001US-0313077P.
XX (CORI-) CORIXA CORP.
PA
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
XX humoral and/or cellular immune response, useful for diagnosing,
XX preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1811; 87dp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
XX associated with cancer, particularly colon cancer. Also disclosed are (i)
XX an expression vector comprising the polynucleotide, (ii) a host cell
XX transformed or transfected with the expression vector, (iii) an isolated
XX antibody, or its antigen-binding fragment, which specifically binds to
XX the polypeptide, (iv) a method of detecting or determining the presence
XX of cancer in a patient, (v) a fusion protein comprising at least one of
XX the polypeptides, (vi) an oligonucleotide that hybridizes to the
XX polynucleotide sequence under highly stringent conditions, and (vii) a
XX method of stimulating and/or expanding T cells specific for a tumour
XX protein. The polypeptide specifically comprises the amino acid sequence
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
XX frames (ORFs) of C636S. These polypeptides are encoded by the
XX polynucleotide sequences, where both are capable of eliciting a humoral
XX and/or cellular immune response. The polynucleotides, polypeptides, and
XX antibodies are useful for diagnosing, preventing or treating cancer,
XX particularly colon cancer. The polynucleotide and polypeptide sequences
XX are also useful in DNA strand invasion, antisense inhibition, mutational
XX analysis, nucleic acid purification, isolation of transcriptionally
XX active genes, blocking or transcription factor binding, genome cleavage
XX or in situ hybridization, and as enhancers of transcription or
XX biomarkers. This sequence represents a human colon cancer associated
XX cDNA. Note: The sequence data for this patent was obtained in electronic
XX format directly from the USPTO web site at seqdata.uspto.gov
XX
XX Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;
SQ
Query Match 90.4%; Score 30; DB 11; Length 552;
Best Local Similarity 64.3%; Pred. No. 0.0026;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYCGTGGCCNTCNCNTAARGAYTAYGARGTNAAR 42
DB 154 GCCACGCTTGGCCGAGCCCTTACAAAGACTATAGAGTAAGA 113
RESULT 5
ADX42774/C
ID ADX42774 standard; cDNA; 552 BP.
XX
XX ADX42774;
AC
XX
XX 21-APR-2005 (first entry)
DT
XX
XX Human cDNA encoding colon cancer protein SEQ ID NO 1811.
DE
XX
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
XX ss; gene.
XX
XX Homo sapiens.
OS
XX
XX WO200274156-A2.
PN
XX
XX 26-SEP-2002.
PD
XX
XX 01-FEB-2002; 2002WO-US002870.
PF
XX
XX 02-FEB-2001; 2001US-0267400P.
PR
11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
PA
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
XX humoral and/or cellular immune response, useful for diagnosing,
XX preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1811; 244pp; English.
XX
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
XX polypeptides, antibodies are useful for diagnosing, preventing or
XX treating cancer, particularly colon cancer. The nucleic acid and
XX polypeptides are also useful in DNA strand invasion, antisense
XX inhibition, mutational analysis, nucleic acid purification, isolation of
XX transcriptionally active genes, blocking or transcription factor binding,
XX genome cleavage or in situ hybridization, and as enhancers of
XX transcription or biomarkers. The kits are useful for detecting antibody
XX binding. The present sequence represents a human cDNA encoding a colon
XX cancer protein.
XX
XX Sequence 552 BP; 126 A; 141 C; 148 G; 137 T; 0 U; 0 Other;
SQ
Query Match 90.4%; Score 30; DB 11; Length 552;
Best Local Similarity 64.3%; Pred. No. 0.0026;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYCGTGGCCNTCNCNTAARGAYTAYGARGTNAAR 42
DB 154 GCCACGCTTGGCCGAGCCCTTACAAAGACTATAGAGTAAGA 113
RESULT 6
AAZ90046
ID AAZ90046 standard; cDNA; 603 BP.
XX
XX AAZ90046;
AC
XX
XX 09-MAY-2000 (first entry)
DT
XX
XX Hydrophobic domain containing protein clone HP10549 coding sequence.
DE
XX
XX Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
XX cell proliferation; immune stimulant; immune deficiency; tumour; pain;
XX rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
XX myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
XX anti-inflammatory; infection; bodily characteristic; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200000506-A2.
PN
XX
XX 06-JAN-2000.
PD
XX
XX 18-JUN-1999; 99WO-JP003242.
PF
XX
XX 26-JUN-1998; 98JP-00180008.
PR
XX
XX (SAGA ) SAGAMI CHEM RES CENT.
PA
XX (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
XX
XX WPI; 2000-160665/14.
XX
XX P-PSDB; AAY78809.
DR
```

XX Novel human proteins having hydrophobic domains used for research and
PT diagnostic purposes.
XX
XX Claim 3; Page 88; 117pp; English.
XX This sequence represents the hydrophobic domain containing protein, clone
CC HP10549 coding region. The sequence is isolated from a human stomach
CC cancer cell line. The invention relates to human proteins with
CC hydrophobic domains, the DNA and the cDNA encoding them. The
CC polynucleotides and proteins are predicted to have biological activities
CC which make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals. Suggested activities include
CC nutritional activity (nutritional source or supplement); cytokine and
CC cell proliferation/differentiation activity; immune stimulating (e.g. as
CC vaccines) or suppressing activity (e.g. to treat various immune
CC deficiencies such as AIDS or HIV, connective tissue disease, systemic
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC organ transplantation); haematopoiesis regulating activity (e.g. in
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or cardiac cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
SQ Sequence 603 BP; 133 A; 179 C; 137 G; 154 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 3; Length 603;
Best Local Similarity 64.3%; Pred. No. 0.0026;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCNCNTGCCNTCCNTAYARGAYTAYGARGTNAAR 42
Db 550 GCCACGCTTGGCGCGCCCTTACAAAGACTATGAGTAAG 591

RESULT 7
ABQ58527/c
ID ABQ58527 standard; cDNA; 620 BP.
AC ABQ58527;
XX
XX 02-AUG-2002 (first entry)
XX Human colon cancer related nucleotide sequence SEQ ID NO:2222.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX Homo sapiens.
XX WO200229086-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US030732.
XX
XX 02-OCT-2000; 2000US-0237271P.
XX (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;

PI Thiaglingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
XX Claim 1; Fig 1; 796pp; English.
XX
XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 620 BP; 152 A; 154 C; 161 G; 142 T; 0 U; 11 Other;
Query Match 90.4%; Score 30; DB 6; Length 620;
Best Local Similarity 64.3%; Pred. No. 0.0027;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAYCNCNTGCCNTCCNTAYARGAYTAYGARGTNAAR 42
Db 154 GCCACGCTTGGCGCGCCCTTACAAAGACTATGAGTAAG 113

RESULT 8
ABQ59698/c
ID ABQ59698 standard; cDNA; 634 BP.
XX
XX ABQ59698;
XX
XX 02-AUG-2002 (first entry)
XX
XX Human colon cancer related nucleotide sequence SEQ ID NO:3393.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX genetic analysis; diagnostic; antisense therapy; gene; ss.
XX Homo sapiens.
XX WO200229086-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US030732.
XX
XX 02-OCT-2000; 2000US-0237271P.
XX (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridizes to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX Sequence 634 BP: 153 A: 154 C: 168 G: 146 T: 0 U: 13 Others:

Query Match 90.4%; Score 30; DB 6; Length 634;
Best Local Similarity 64.3%; Pred. No. 0.0627;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCNTGGCCNTCCNCTAYAAARGAYTAYARGGTNAAR 42
158 GCCCACGCTTGGCCGAGCGCTTACAAAGACTATGAAGTAAAG 117

Db

RESULT 9	
ABT22328/c	
ID	ABT22328 standard; DNA; 642 BP.
XX	
XX	
AC	ABT22328;
XX	
XX	
DT	16-APR-2003 (first entry)
XX	
XX	
DE	Breast cancer marker gene SEQ ID NO 701.

KW	Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;
KW	chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;
KW	surrogate marker gene; pharmacodynamic marker gene; transgenic animal;
KW	human; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200285298-A2.
XX	
XX	
PD	31-OCT-2002.
XX	
PF	19-APR-2002; 2002WO-US012612.
XX	
PR	20-APR-2001; 2001US-0285163P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;
XX	
DR	WPI: 2003-093053/08.

Novel isolated polypeptide encoded by breast cancer marker gene, useful for diagnosing, staging, monitoring, prognosing and treating diseases associated with breast cancer.

Disclosure; Page 187-188; 725pp; English.

The invention relates to an isolated polypeptide encoded by a breast cancer marker gene comprising any of 1417 21-805 nucleotide sequences, given in the specification. The methods of the invention are useful for diagnosing patients having an identified breast mass or symptoms associated with breast cancer, to diagnose breast cancer or its precursors, and for monitoring the efficacy of treatment of a breast cancer patient (e.g. efficacy of chemotherapy). The methods are also

useful for evaluating a patient before, after or during therapy, to evaluate the reduction in a tumour burden. The breast cancer marker gene proteins are useful as immunogens for raising antibodies, by immunising a mammal with a breast cancer marker protein. The marker proteins are useful as bait proteins in a two-hybrid or three-hybrid assay, to identify other proteins which bind to or interact with the marker proteins. The breast cancer marker genes are useful as surrogate marker genes for one or more disorders, disease states or conditions leading to disease states, in particular, breast cancers. The breast cancer marker genes are useful as pharmacodynamic marker genes. An antibody which selectively binds to a protein of a breast cancer marker gene is useful for treating cancers, particularly breast cancers. The host cell of the invention is useful for producing non-human transgenic animals. This polynucleotide sequence represents one of the breast cancer marker genes of the invention

Sequence 642 BP: 148 A; 160 C; 180 G; 142 T; 0 U; 12 Other;

Query Match	90.4%	Score 30;	DB 10;	Length 642;
Best Local Similarity	64.3%	Pred. No. 0.0027;		
Matches 27;	Conservative	7;	Mismatches 8;	Indels 0;
Gaps 0;				
Qy	1	GCNCAYGCNTGGCCNTCCNCTAYAAARGAYTAYGARGCTNAAR	42	
Db	184	GCCACGCTTTGGCCGAGCCCTTACAAGAGCTATGAAGTAAAG	143	

RESULT 10	
ADO30035	
ID	ADO30035 standard; cDNA; 1074 BP.
XX	
AC	
XX	ADO30035;
DT	29-JUL-2004 (first entry)
XX	
DE	Human GPCR RAI3 polynucleotide, SEQ ID NO:1137.
XX	
KW	G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW	transgenic mouse; neurological disorder; adrenal gland disorder;
KW	colon disorder; intestinal disorder; cardiovascular disorder;
KW	muscular disorder; blood disorder; immune disorder; bone disorder;
KW	joint disorder; metabolic disorder; nutritive disorder; cancer;
KW	kidney disorder; liver disorder; lung disorder; breast disorder;
KW	ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW	skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW	thyroid disorder; thyroid disorder; antiparkinsonian; antimanic;
KW	cystatic; antiinflammatory; vasotropic; antitanginal; antiarrhythmic;
KW	CNS; central nervous system; respiratory; anti diarrhoeic; antidiabetic;
KW	virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW	dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW	immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW	gene; ss.

OS	Homo sapiens.
XX	
FN	WO2004040000-A2.
XX	
XX	
PD	13-MAY-2004.
XX	
XX	
PF	09-SEP-2003; 2003WO-US028226.
XX	
PR	09-SEP-2002; 2002US-0409303P.
PR	09-APR-2003; 2003US-0461329P.
XX	
PA	(PRIM-) PRIMAL INC.

XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li P;
 EI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
 XX WPI; 2004-390329/36.
 DR P-PSDB; ADO29632.
 XX Novel mammalian G protein coupled receptors, useful for identifying

PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
PS Claim 151; SEQ ID NO 1137; 542pp; English.
XX
CC The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1074 BP; 215 A; 316 C; 259 G; 284 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 12; Length 1074;
Best Local Similarity 64.3%; Pred. No. 0.0031;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
DB 1018 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1059

RESULT 11
AA261776
ID AA261776 standard; cDNA; 1212 BP.
AC
AC AA261776;
XX
XX 27-MAR-2000 (first entry)
DE cDNA encoding human skin cell transmembrane protein, SEQ ID NO:249.
XX
XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumour vascularisation; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS Homo sapiens.
XX
XX WO955865-A1.
PN
XX 04-NOV-1999.
PD
XX 29-APR-1999; 99WO-NZ000051.
PF

XX
PR 29-APR-1998; 98US-00069726.
PR 09-NOV-1998; 98US-00188930.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA
XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
PI WPI; 2000-072177/06.
DR
XX Novel polynucleotides useful for the treatment of various conditions
PT including wounds and cancer.
PT
XX Claim 1; Page 155; 235pp; English.
PS
XX The invention relates to novel nucleic acid sequences derived from rat
CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC cells. Polypeptides of the invention may be used to treat inflammation,
CC cancer and neurological diseases. The proteins may be used to stimulate
CC the growth and motility of keratinocytes, to inhibit the growth of cancer
CC cells, to modulate angiogenesis and tumour vascularisation, to modulate
CC skin inflammation, to modulate epithelial cell growth and to inhibit
CC binding of HIV-1 to leukocytes. The invention may also be used to treat
CC growth and developmental defects, skin wounds and hair follicle
CC disorders. Sequences AA261606-261832 represent cDNA sequences derived
CC from several mouse, rat or human skin cell types. Sequences AA261606-
CC 261649, AA261725-261765, AA261802-261811 and AA261826 encode proteins
CC with an N-terminal signal sequence, indicating that the proteins are
CC secreted. Sequences AA261650-261668, AA261766-261780, AA261812-261817 and
CC AA261827-261829 encode proteins with one or more putative transmembrane
CC domains
XX
SQ Sequence 1212 BP; 242 A; 350 C; 312 G; 308 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 3; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.0032;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
DB 990 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 12
AAC99709
ID AAC99709 standard; cDNA; 1212 BP.
XX
AC AAC99709;
XX
XX 08-MAR-2001 (first entry)
DE Skin cell cDNA, SEQ ID NO: 249.
XX
XX Human; skin cell; cytostatic; antiinflammatory; anti-HIV; neurotropic;
KW neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200069884-A2.
PN
XX 23-NOV-2000.
PD
XX 15-MAY-2000; 2000WO-NZ000075.
PF
XX 14-MAY-1999; 99US-00312283.
PR
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA
XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;
PI Murison JG;


```

DR WPI; 2005-371577/38.
DR P-PSDB; AEA00710.
XX
PT Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds
PT to tumor-associated antigenic target polypeptide, useful for diagnosing
PT or treating cancer.
XX
PS Claim 1; SEQ ID NO 64; 96pp; English.
XX
CC This invention relates to a novel isolated antibody, for example anti-E16
CC or anti-TAT112 antibody, that binds to a tumor-associated antigenic
CC target polypeptide (TAT) and that lacks an associated signal peptide
CC sequence. The invention may be useful for the development of compounds
CC with a cytostatic activity acting as antagonists of the TAT188
CC polypeptide or RNA interference whilst the disclosed sequences may be
CC useful for gene therapy. The invention is useful for inducing the death
CC of a cell (such as a cancer cell chosen from breast, colon, rectum,
CC endometrium, kidney, lung, ovary, skin and liver) to which it binds,
CC inhibiting proliferation or promoting cell death of a cell expressing
CC TAT188. In addition, the invention may be useful for detecting the level
CC of TAT188 polypeptide in a test cell relative to a control cell, or
CC treating or preventing a cell proliferative disorder associated with
CC increased expression of TAT188. The novel antibody of the invention is
CC useful for inhibiting the growth of a cancer cell and may be useful for
CC diagnosing or treating cancer. The present sequence is that of the human
CC TAT143 cDNA which encodes a protein against which an antibody of the
CC invention may be targeted.
XX
SQ Sequence 1460 BP; 302 A; 418 C; 362 G; 378 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 14; Length 1460;
Best Local Similarity 64.3%; Pred. No. 0.0034;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCVTGGCCGTCNCNTAYARGAYTAYGAGTNAAR 42
|||:||||| |||:|||||:|||||:|:
Db 1145 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1186

RESULT 17
AAF58615
ID AAF58615 standard; cDNA; 1619 BP.
XX
AC AAF58615;
XX
DT 24-APR-2001 (first entry)
XX
DE Human RECAP polynucleotide, SEQ ID NO: 43.
XX
KW Human; RECAP; receptors and associated proteins; cerebroprotective;
KW neurotropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW cytostatic; antibacterial; virucide; fungicide; protozoacide;
KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200107612-A2.
XX
PD 01-FEB-2001.
XX
PF 21-JUL-2000; 2000WO-US020035.
XX
PR 21-JUL-1999; 95US-0145232P.
PR 07-OCT-1999; 95US-0158578P.
PR 12-NOV-1999; 99US-0165192P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
PI Baughn MR, Lu DM, Hillman JL, Patterson C, Lal P;
XX

```

```

DR WPI; 2001-168554/17.
DR P-PSDB; AAB68891.
XX
PT Novel receptors and associated proteins for diagnosis and treatment of
PT neurological disorders, immunological disorders including autoimmune/
PT inflammatory disorders and cell proliferative disorders such as cancer.
XX
PS Example 5; Page 126-127; 128pp; English.
XX
CC The present sequence encodes a human RECAP (receptors and associated
CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
CC in the diagnosis, treatment and prevention of neurological disorders such
CC as stroke, Alzheimer's disease, Pick's disease, Huntington's disease,
CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral
CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD
CC (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker
CC syndrome); immunological disorders, including autoimmune/inflammatory
CC disorders such as AIDS, Digeorge's syndrome, severe combined
CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,
CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis and cancer
XX
SQ Sequence 1619 BP; 331 A; 463 C; 400 G; 425 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 4; Length 1619;
Best Local Similarity 64.3%; Pred. No. 0.0035;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCVTGGCCGTCNCNTAYARGAYTAYGAGTNAAR 42
|||:||||| |||:|||||:|||||:|:
Db 1137 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1178

RESULT 18
AAZ90056
ID AAZ90056 standard; cDNA; 1718 BP.
XX
AC AAZ90056;
XX
DT 09-MAY-2000 (first entry)
XX
DE Hydrophobic domain containing protein clone HP10549 nucleotide sequence.
XX
KW Hydrophobic domain; clone HP10549; nutritional supplement; SCID; HIV;
KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW anti-inflammatory; infection; bodily characteristic; ss.
XX
OS Homo sapiens.
XX
PN WO200000506-A2.
XX
PD 06-JAN-2000.
XX
PF 18-JUN-1999; 99WO-JP003242.
XX
PR 26-JUN-1998; 98JP-00180008.
XX
PA (SAGA ) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
PI Kato S, Kimura T;
XX
WPI; 2000-160665/14.
DR P-PSDB; AAY78809.
XX
PT Novel human proteins having hydrophobic domains used for research and
PT diagnostic purposes.

```

XX PS Claim 4; Page 111-113; 117pp; English.

XX CC This sequence represents the hydrophobic domain containing protein, clone

XX CC HP10549 nucleotide sequence. The sequence is isolated from a human

XX CC stomach cancer cell line. The invention relates to human proteins with

XX CC hydrophobic domains, the DNA and the cDNA encoding them. The

XX CC polynucleotides and proteins are predicted to have biological activities

XX CC which make them suitable for treating, preventing or ameliorating medical

XX CC conditions in humans and animals. Suggested activities include

XX CC nutritional activity (nutritional source or supplement); cytokine and

XX CC cell proliferation/differentiation activity; immune stimulating (e.g. as

XX CC vaccines) or suppressing activity (e.g. to treat various immune

XX CC deficiencies such as AIDS or HIV, connective tissue disease, systemic

XX CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary

XX CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin

XX CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease

XX CC and autoimmune inflammatory eye disease, as well as asthma, allergies and

XX CC organ transplantation); haematopoiesis regulating activity (e.g. in

XX CC treatment of myeloid or lymphoid cell deficiencies); tissue growth

XX CC activity (e.g. wound healing and tissue repair, ulcers, burns,

XX CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic

XX CC activity; haemostatic and thrombolytic activity (e.g. treating

XX CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and

XX CC tumour inhibition activity. The polynucleotides are also stated to be

XX CC useful for gene therapy. Other activities include inhibiting infections

XX CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,

XX CC malaria); affecting bodily characteristics such as, e.g. weight, colour,

XX CC skin, affecting biorhythms or circadian cycles; enhancing fertility;

XX CC treatment of depression; treatment of pain; hormonal or endocrine

XX CC activity. The polynucleotides may also be used for recombinant expression

XX CC of the protein

XX CC

XX SQ Sequence 1718 BP; 381 A; 495 C; 377 G; 465 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 3; Length 1718;

Best Local Similarity 64.3%; Pred. No. 0.0035;

Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Oy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42

Db 561 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 602

RESULT 19

ADF70573

ID ADF70573 standard; DNA; 1788 BP.

XX AC

XX ADP70573;

XX DT

XX DT 12-FEB-2004 (first entry)

XX DE Orphan receptor ligand-related human protein gene SeqID196.

XX DE ligand; orphan receptor protein; fusion protein; fluorescent protein;

XX KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;

XX KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.

XX OS Homo sapiens.

XX PN WO2003071272-A1.

XX PD

XX PD 28-AUG-2003.

XX PF

XX PF 21-FEB-2003; 2003WO-JP001901.

XX PR

XX PR 22-FEB-2002; 2002JP-00045728.

XX PR 23-JUL-2002; 2002JP-00213949.

XX PR 11-OCT-2002; 2002JP-00298237.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;

XX WPI; 2003-697654/66.

DR P-PSDB; ADF70471.

XX PT Transformation of cells with a fusion protein of an orphan receptor

XX PT protein with a fluorescent protein useful for identification of ligands

XX PT to the orphan receptor.

XX PS Example 4; SEQ ID NO 196; 594pp; Japanese.

XX CC This invention relates to a novel method of identifying ligands to an

XX CC orphan receptor protein which comprises transforming cells with DNA

XX CC encoding a fusion protein of the orphan receptor with a fluorescent

XX CC protein, so that the fusion protein is expressed in the cells (or cell

XX CC membranes isolated from them) and contacting the cells with the potential

XX CC ligand to be tested. A suitable fluorescent protein for incorporation in

XX CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,

XX CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the

XX CC identification of ligands binding to an orphan receptor protein.

XX SQ Sequence 1788 BP; 444 A; 473 C; 412 G; 459 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 10; Length 1788;

Best Local Similarity 64.3%; Pred. No. 0.0036;

Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Oy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42

Db 1018 GCCACGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1059

RESULT 20

ADF70573

ID ADF70573 standard; cDNA; 2302 BP.

XX AC

XX ABT10173;

XX DT

XX DT 04-DEC-2002 (first entry)

XX DE Human breast cancer associated coding sequence SEQ ID NO: 307.

XX KW Human; breast specific gene; breast cancer; differential expression;

XX KW cytostatic; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200259271-A2.

XX PD

XX PD 01-AUG-2002.

XX PF

XX PF 25-JAN-2002; 2002WO-US002176.

XX PR

XX PR 25-JAN-2001; 2001US-0263757P.

XX PR 25-APR-2001; 2001US-0286090P.

XX PR 23-MAY-2001; 2001US-0292517P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Orr MS, Nation M, Diggans JC, Zeng W;

XX WPI; 2002-674803/72.

XX DR

XX PT Diagnosing breast cancer in a patient comprises detecting the level of

XX PT gene expression in cell or tissue samples, where a differential gene

XX PT expression is indicative of breast cancer.

XX PS Claim 1; SEQ ID NO 307; 260pp + Sequence Listing; English.

XX CC The present invention relates to methods of diagnosing breast cancer in a

XX CC patient, which comprise detecting the level of expression in a tissue

XX CC sample of two or more genes selected from those shown in AB09867-

XX CC ABT11112, where a differential expression of the genes indicates breast

XX CC cancer. The methods are useful in diagnosing, treating, detecting the

CC progression, and in monitoring treatment of breast cancer in patients.
CC The methods are also useful as a screening tool for agents that modulate
CC the onset or progression of breast cancer. The breast cancer genes may be
CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub.published_pct_sequences
XX
SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 6; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0038;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYCGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158
DB 1117 GCCCAGCGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 21

ACC58386
ID ACC58386 standard; cDNA; 2302 BP.

XX ACC58386;

XX 26-AUG-2003 (first entry)

XX Human GPCR-like retinoic acid-induced gene 1.

XX Human; retinoic acid-induced gene 1; RAIG1; feeding; fasting; GPCR;
XX receptor; G-protein coupled receptor; anorectic; antidiabetic;
XX antidepressant; immunomodulator; transgenic; gene therapy; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 100..1173
FT /*tag= a
FT /product= "GPCR-like RAIG1"

XX WO2003016553-A2.

XX 27-FEB-2003.

XX 20-AUG-2002; 2002WO-US026510.

XX 20-AUG-2001; 2001US-0313940P.

XX (GETH) GENENTECH INC.
XX (CURA-) CURAGEN CORP.

XX Lewin DA, Stewart TA;

XX WPI; 2003-278580/27.

XX P-PSDB; ABR42649.

XX New G-protein coupled receptor-like retinoic acid induced gene 1 (GPCR-
XX like RAIG1) polypeptide and gene, useful for diagnosing or treating
XX metabolic disorders, e.g. obesity, anorexia, cachexia or diabetes.

XX Claim 10; Page 16-17; 150pp; English.

XX The present sequence is that of human G-protein coupled receptor-like
XX retinoic acid induced gene 1 (GPCR-like RAIG1). This is the human
XX homologue of murine GPCR-like RAIG1. The murine gene was shown to be
XX differentially regulated during fasting-feeding cycles in mice, with
XX moderate induction early in fasting, down-regulation with extended
XX fasting and 4-fold up-regulation with feeding in recovery from fasting.
XX The differentially expressed gene, its mRNA, and the encoded protein, can
XX each be manipulated to detect and treat metabolic disorders associated

CC with up- or down-regulation of GPCR-like RAIG1 activity, such as obesity,
CC anorexia, cachexia or diabetes
XX
SQ Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 8; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0038;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYCGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158
DB 1117 GCCCAGCGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 22

ADD93240
ID ADD93240 standard; cDNA; 2302 BP.

XX ADD93240;

XX 29-JAN-2004 (first entry)

XX RAIG1 coding sequence.

XX ss; gene; retinoic acid-inducible gene 1; RAIG1; orphan;
XX G-protein coupled receptor; GPCR; chromosome 12; GPCR5B; carcinoma;
XX vaccine; breast cancer; pancreatic cancer; lung cancer; liver cancer;
XX ovarian cancer; colon cancer; osteosarcoma.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 100..1173
FT /*tag= a
FT /product= "RAIG1"

XX WO2003087832-A2.

XX 23-OCT-2003.

XX 10-APR-2003; 2003WO-GB001587.

XX 11-APR-2002; 2002GB-00008331.

XX 17-SEP-2002; 2002GB-00021538.

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX Terrett JA;

XX WPI; 2003-845382/78.

XX P-PSDB; ADD93239.

XX GENBANK; AF095448.

XX Screening, diagnosing and/or treating carcinoma, including breast,
XX pancreatic, lung, liver, ovarian and colon cancer by detecting the change
XX in expression or activity of an RAIG1 polypeptide or encoding nucleic
XX acid molecule.

XX Claim 1; Fig 2; 43pp; English.

XX This sequence encodes retinoic acid-inducible gene 1 (RAIG1) polypeptide.
XX RAIG1 is an orphan G-protein coupled receptor (GPCR) located on
XX chromosome 12. RAIG1 shows a restricted expression pattern compared to a
XX related receptor, GPCR5B, which is widely expressed in peripheral and
XX central tissues. The RAIG1 cDNA and polypeptide sequence may be used in
XX the method of the invention for screening for and/or diagnosis of
XX carcinoma in a subject, and/or monitoring the effectiveness of carcinoma
XX therapy. The method comprises detecting and/or quantifying in a
XX biological sample obtained from the subject an RAIG1 polypeptide and a
XX nucleic acid molecule. The RAIG1 polypeptide and nucleic acid molecule
XX are useful in the manufacture of a medicament for the treatment of
XX carcinoma, where the composition is a vaccine. An agent which interacts
XX with or causes change in the expression or activity of an RAIG1

CC polypeptide or nucleic acid molecule, is also useful in the manufacture
CC of a medicament for the treatment of carcinoma that is breast cancer,
CC pancreatic cancer, lung cancer, liver cancer, ovarian cancer, colon
CC cancer and/or osteosarcoma. They can also be used in the diagnosis and
CC screening of such carcinomas.

XX Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 10; Length 2302;

Best Local Similarity 64.3%; Pred. No. 0.0038;

Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNGTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:|||||

Db 1117 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 23

ADL24773

ID ADL24773 standard; DNA; 2302 BP.

XX AC ADL24773;

XX 20-MAY-2004 (first entry)

XX Intestinal epithelium/peyer's patch M cell-associated DNA sequence #171.

XX intestinal epithelium cell development; peyer's patch M cell development;
KW inflammatory bowel disease; glutenenteropathy; infectious disease;
KW autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;
KW Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;
KW immune system disorder; hypersensitivity; anaphylaxis;
KW blood group incompatibility; ds; human.

XX Homo sapiens.

XX WO200280852-A2.

XX 17-OCT-2002.

XX 04-APR-2002; 2002WO-US010873.

XX 04-APR-2001; 2001US-0281416P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;

XX WPI; 2003-075470/07.

XX Novel isolated or purified polypeptide encoded by genes associated with
PT intestinal epithelium or M cell development, differentiation or function,
PT useful for treating autoimmune diseases and infectious diseases.

XX Claim 1; SEQ ID NO 283; 152pp; English.

XX The invention comprises DNA sequences which are associated with
CC intestinal epithelium and peyer's patch M cells. The DNA sequences of the
CC invention are useful for assessing, modifying, modulating or regulating
CC intestinal epithelium or M cell development. The DNA sequences of the
CC invention are also useful in the treatment of: inflammatory bowel
CC disease, glutenenteropathy, infectious diseases, autoimmune diseases
CC (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's
CC disease, multiple sclerosis, allergy, asthma and diabetic mellitus),
CC diseases or disorders of the immune system, hypersensitivity,
CC anaphylaxis, and blood group incompatibility. The present nucleic acid
CC represents an intestinal epithelium/peyer's patch M cell-associated DNA
CC sequence of the invention. NOTE: The present sequence is not shown in the
CC specification, but has been retrieved from the WIPO website.

XX Sequence 2302 BP; 494 A; 666 C; 533 G; 609 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 10; Length 2302;

Best Local Similarity 64.3%; Pred. No. 0.0038;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNGTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:|||||

Db 1117 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 24

ABT31923

ID ABT31923 standard; DNA; 2316 BP.

XX AC ABT31923;

XX 01-MAY-2003 (first entry)

XX Human breast cancer / ovarian cancer related coding sequence #30.

XX Human; gene; ds; cytostatic; breast cancer; ovarian cancer.

XX Homo sapiens.

XX WO2003000012-A2.

XX 03-JAN-2003.

XX 21-JUN-2002; 2002WO-US019773.

XX 21-JUN-2001; 2001US-0300159P.

XX 27-JUN-2001; 2001US-0301351P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Veiby OP;

XX WPI; 2003-267848/26.

XX P-PSDB; ABJ37054.

XX Determining the presence of breast cancer in an individual, involves
PT using specific polynucleotide markers.

XX Disclosure; Page 163; 233pp; English.

XX The invention comprises a method for assessing whether a patient is
CC afflicted with breast cancer or ovarian cancer. The method involves the
CC use of specific DNA markers. The method of the invention is useful in the
CC detection and treatment of ovarian and breast cancer. DNA sequences
CC ABT31894 - ABT31949 encode human breast/ovarian cancer-related proteins

XX Sequence 2316 BP; 502 A; 670 C; 535 G; 609 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 10; Length 2316;

Best Local Similarity 64.3%; Pred. No. 0.0039;

Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNGTGGCCNTCCNTAYAARGAYTAYGARGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:|||||

Db 1123 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1164

RESULT 25

AAH14688

ID AAH14688 standard; cDNA; 2446 BP.

XX AC AAH14688;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:12388.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

```
XX EP1074617-A2.
FN
XX
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 12388; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 2446 BP; 507 A; 723 C; 572 G; 644 T; 0 U; 0 Other;
SQ
Query Match 90.4%; Score 30; DB 4; Length 2446;
Best Local Similarity 64.3%; Pred. No. 0.0039;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYCGNTGGCCNTCCNTAYARGAYTAYGAGCTGTAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 26
ABZ42832
ID ABZ42832 standard; DNA; 2456 BP.
XX
XX ABZ42832;
XX
XX 04-MAR-2003 (first entry)
XX
XX Human G protein-coupled receptor RAIG1 nucleotide SEQ ID NO:453.
XX
XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
XX G protein-coupled receptor modulator; antibody; immune-related disease;
```

growth-related disease; cell regeneration-related disease; AIDS; cancer; immunological-related cell proliferative disease; autoimmune disease; Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy; osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes; graft versus host disease; Parkinson's disease; multiple sclerosis; pain; psoriasis; anxiety; depression; schizophrenia; dementia; memory loss; mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea; hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma; ulcer; gene; ds.

Homo sapiens.

WO200261087-A2.

08-AUG-2002.

19-DEC-2001; 2001WO-US050107.

19-DEC-2000; 2000US-0257144P.

(LIFE-) LIFESPAN BIOSCIENCES INC.

Burmer GC, Roush CL, Brown JP;

WPI; 2003-046718/04.

P-PSDB; ABP81984.

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention

Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;

Query Match 90.4%; Score 30; DB 8; Length 2456;

Best Local Similarity 64.3%; Pred. No. 0.0039;

Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYCGNTGGCCNTCCNTAYARGAYTAYGAGCTGTAAR 42

Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 27

ACCT2695

ID ACCT72695 standard; cDNA; 2456 BP.
XX ACC72695;
XX
XX 09-JUL-2003 (first entry)
XX
XX Human cancer related protein encoding cDNA SEQ ID NO:34.
XX
XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometriosis; gene; ss.
XX Homo sapiens.
XX WO2003025138-A2.
XX
XX 27-MAR-2003.
XX
XX 17-SEP-2002; 2002WO-US029560.
XX
XX 17-SEP-2001; 2001US-0323469P.
XX 20-SEP-2001; 2001US-0323887P.
XX 13-NOV-2001; 2001US-0350666P.
XX 08-FEB-2002; 2002US-0355145P.
XX 08-FEB-2002; 2002US-0355257P.
XX 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI Zlotnik A;
XX
XX WPI; 2003-354600/33.
XX P-PSDB; ABR58568.
XX
XX New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischaemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
XX Claim 8; Page 643-644; 767pp; English.
XX
XX The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SQ Sequence 2456 BP; 522 A; 720 C; 571 G; 643 T; 0 U; 0 Other;
Query Match 90.4%; Score 30; DB 10; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0039;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GCNCAVCGTCCNTCNCNTAARGATYAGTGNTAAR 42
Db 1271 GCCCAGCGTTGGCCGAGCCCTTACAAGACTGATGAAGTAAAG 1312
RESULT 28

ADN39301
ID ADN39301 standard; cDNA; 2456 BP.
XX
XX AC ADN39301;
XX
XX 17-JUN-2004 (first entry)
XX
XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:619.
XX
XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
XX
XX 13-NOV-2001; 2001US-0350666P.
XX 21-NOV-2001; 2001US-0322464P.
XX 29-NOV-2001; 2001US-0334393P.
XX 03-DEC-2001; 2001US-0335394P.
XX 14-DEC-2001; 2001US-0340376P.
XX 08-JAN-2002; 2002US-0347211P.
XX 10-JAN-2002; 2002US-0347349P.
XX 08-FEB-2002; 2002US-0355250P.
XX 13-FEB-2002; 2002US-0356714P.
XX 20-FEB-2002; 2002US-0359077P.
XX 29-MAR-2002; 2002US-0368809P.
XX 04-APR-2002; 2002US-0370110P.
XX 12-APR-2002; 2002US-0372246P.
XX 05-JUN-2002; 2002US-0386614P.
XX 16-JUL-2002; 2002US-0396839P.
XX 22-JUL-2002; 2002US-0397757P.
XX 22-JUL-2002; 2002US-0397845P.
XX 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
XX P-PSDB; ADN39302.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 619; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may


```
FT CDS 254..1327
FT /*tag= a
FT /product= "TAT115"
XX
XX
XX US2005106644-A1.
XX
XX 19-MAY-2005.
XX
XX 08-SEP-2004; 2004US-00936626.
XX
XX 20-JUN-2001; 2001US-0299500P.
XX
XX 29-JUN-2001; 2001US-0301880P.
XX
XX 18-SEP-2001; 2001US-0323288P.
XX
XX 19-JUN-2002; 2002US-00177488.
XX
XX 26-MAR-2004; 2004US-0557116P.
XX
XX 04-AUG-2004; 2004US-0598899P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
XX Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
XX Sliwkowski M;
XX
XX WPI; 2005-384304/39.
XX P-PSDB; AEA00166.
XX
XX Novel isolated antibody capable of binding to tumor-associated antigenic
XX target polypeptide, useful for treating cell proliferative disorder e.g.
XX cancer.
XX
XX Claim 1; SEQ ID NO 40; 337pp; English.
XX
XX The invention relates to a novel isolated antibody binding to a
XX polypeptide having at least 80% sequence identity to a polypeptide having
XX any one of 76 fully defined 182-910 amino acid tumor-associated antigenic
XX target polypeptide (TAT) sequences (AEA00127-AEA00202) given in the
XX specification, a polypeptide having any one of (AEA00127-AEA00202),
XX lacking its associated signal peptide, or an extracellular domain of a
XX polypeptide having any one of (AEA00127-AEA00202). The polypeptide is
XX encoded by the nucleotide sequence having any one of (AEA00049-AEA00126).
XX An antibody of the invention has cytostatic activity. The antibody is
XX useful for inhibiting growth of a cell expressing TAT188, which involves
XX contacting the cell with the antibody. The cell is a cancer cell chosen
XX from breast, colon, rectum, endometrium, kidney, lung, ovary, skin and
XX liver cell. The cancer cell is a mammalian cell, preferably a human cell.
XX The antibody is also useful for detecting the level of TAT188
XX polypeptide expressed in a test cell relative to a control cell, and for
XX detecting the level of TAT188 polypeptide or a polypeptide having at
XX least 80 % sequence identity to the TAT188 polypeptide sequence in a test
XX cell relative to a control cell. The antibody is useful for treating a
XX preventing a cell proliferative disorder associated with increased
XX expression or activity of a polypeptide having at least 80 % identity to
XX a TAT188 polypeptide sequence. The cell proliferative disorder is cancer.
XX The method of the invention is useful for inhibiting the growth of a
XX cancer cell. The present sequence encodes a polypeptide of the invention.
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
XX
XX Query Match 90.4%; Score 30; DB 14; Length 2456;
XX Best Local Similarity 64.3%; Pred. No. 0.0039;
XX Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
XX
XX 1 GCNCAYCGNTGGCCNTCCNCTAAYARGAYTAYGARGTNAAR 42
XX |||||
XX 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
XX
XX RESULT 36
XX AEA00608
XX ID AEA00608 standard; cDNA; 2456 BP.
XX
XX AC AEA00608;
XX
```

```
DT 28-JUL-2005 (first entry)
XX
XX Human TAT115 cDNA sequence SeqID40.
XX
XX antibody identification; tumor-associated antigen; cytostatic;
XX RNA interference; gene therapy; cell death; cancer; breast tumor;
XX colon tumor; rectal tumor; renal tumor; lung tumor; ovary tumor;
XX skin tumor; liver tumor; gene; ss; TAT115.
XX
XX Homo sapiens.
XX
XX US2005107595-A1.
XX
XX 19-MAY-2005.
XX
XX 10-SEP-2004; 2004US-00938061.
XX
XX 20-JUN-2001; 2001US-0299500P.
XX
XX 29-JUN-2001; 2001US-0301880P.
XX
XX 18-SEP-2001; 2001US-0323288P.
XX
XX 19-JUN-2002; 2002US-00177488.
XX
XX 26-MAR-2004; 2004US-0557116P.
XX
XX 04-AUG-2004; 2004US-0598899P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Cairns B, Chen R, Frantz G, Hillan KJ, Koepfen H, Phillips HS;
XX Polakis P, Spencer SD, Smith V, Williams PM, Wu TD, Zhang Z;
XX Sakanaka C, Chuntharapai A, Reed CJ;
XX
XX WPI; 2005-371577/38.
XX P-PSDB; AEA00686.
XX
XX Novel isolated antibody e.g. anti-E16 or anti-TAT112 antibody that binds
XX to tumor-associated antigenic target polypeptide, useful for diagnosing
XX or treating cancer.
XX
XX Claim 1; SEQ ID NO 40; 96pp; English.
XX
XX This invention relates to a novel isolated antibody, for example anti-E16
XX or anti-TAT112 antibody, that binds to a tumor-associated antigenic
XX target polypeptide (TAT) and that lacks an associated signal peptide
XX sequence. The invention may be useful for the development of compounds
XX with a cytostatic activity acting as antagonists of the TAT188
XX polypeptide or RNA interference whilst the disclosed sequences may be
XX useful for gene therapy. The invention is useful for inducing the death
XX of a cell (such as a cancer cell chosen from breast, colon, rectum,
XX endometrium, kidney, lung, ovary, skin and liver) to which it binds,
XX inhibiting proliferation or promoting cell death of a cell expressing
XX TAT188. In addition, the invention may be useful for detecting the level
XX of TAT188 polypeptide in a test cell relative to a control cell, or
XX treating or preventing a cell proliferative disorder associated with
XX increased expression of TAT188. The novel antibody of the invention is
XX useful for inhibiting the growth of a cancer cell and may be useful for
XX diagnosing or treating cancer. The present sequence is that of the human
XX TAT115 cDNA which encodes a protein against which an antibody of the
XX invention may be targeted.
XX
XX Sequence 2456 BP; 519 A; 722 C; 571 G; 644 T; 0 U; 0 Other;
XX
XX Query Match 90.4%; Score 30; DB 14; Length 2456;
XX Best Local Similarity 64.3%; Pred. No. 0.0039;
XX Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
XX
XX 1 GCNCAYCGNTGGCCNTCCNCTAAYARGAYTAYGARGTNAAR 42
XX |||||
XX 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312
XX
XX RESULT 37
XX AED47480
XX ID AED47480 standard; DNA; 2456 BP.
XX
```


AC ADX97494;
XX 21-APR-2005 (first entry)
XX Pancreatic cancer associated human protein encoding DNA, SEQ ID 42.
XX pancreas tumor; cytostatic; gene; ds.
XX Homo sapiens.
XX EP1471075-A2.
XX 27-OCT-2004.
XX 31-MAR-2004; 2004EP-00090124.
XX 31-MAR-2003; 2003DE-01015834.
XX (HINZ/) HINZMANN B.
XX (ROSE/) ROSENTHAL A.
XX (PILA/) PILARSKY C.
XX (DAHL/) DAHL E.
XX (SPEC/) SPECHT T.
XX (LICH/) LICHTNER R.
XX Rosenthal A, Pilarsky C, Dahl E, Specht T, Bruemendorf T;
XX Lichtner R, Staub E, Roepcke S, Li X;
XX WPI; 2004-768082/76.
XX P-PSDB; ADX97565.
XX New nucleic acid differentially expressed in pancreatic tumor tissue, for
XX use as diagnostic agents and in screening for therapeutic agents.
XX Claim 1; SEQ ID NO 42; 28pp; German.
XX The invention relates to a novel human nucleic acid sequence of the
XX pancreas and its encoded protein. The invention further comprises:
XX proteins and peptides, preferably isolated, that contain a sequence
XX encoded by the novel nucleic acid; and methods for diagnosis and
XX treatment of pancreatic cancer, using a substance that inhibits or binds
XX to the protein or its DNA, including: an antisense oligonucleotide, short
XX interfering RNA or ribozyme directed against the pancreatic protein, an
XX organic molecule, particularly having a molecular weight below 5000,
XX especially 300, that binds to the pancreatic DNA, an aptamer or
XX (monoclonal) antibody, preferably human or humanized, that binds to the
XX pancreatic DNA, or an anti-idiotypic antibody raised against the
XX monoclonal antibody, any of which may be derivatized with a reporter
XX group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
XX human pancreatic proteins and their encoding DNA have cytostatic
XX activity. The novel sequences are useful for inhibiting transcription
XX and/or expression of genes and proteins associated with pancreatic
XX cancer. This polynucleotide sequence represents the DNA encoding one of
XX the novel human pancreatic proteins of the invention. Note: This sequence
XX is not shown in the specification, it has been electronically downloaded
XX from a DVD-rom provided with this specification by the European Patent
XX Office.
XX
XX Sequence 6730 BP; 1469 A; 1763 C; 1555 G; 1943 T; 0 U; 0 Other;
SQ Query Match 90.4%; Score 30; DB 13; Length 6730;
Best Local Similarity 64.3%; Pred. No. 0.0052;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYGCNTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
DB 1271 GCCCAGCGTTGGCCGACGCTTACAAAGACTATGAGTAAG 1312
RESULT 42
AAH08354
ID AAH08354 standard; cDNA; 680 BP.
XX

AC AAH08354;
XX 26-JUN-2001 (first entry)
XX Human cDNA clone (5'-primer) SEQ ID NO:5189.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX Claim 1; SEQ ID NO 5189; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 680 BP; 143 A; 194 C; 151 G; 184 T; 0 U; 8 Other;
SQ Query Match 82.5%; Score 27.4; DB 4; Length 680;
Best Local Similarity 59.5%; Pred. No. 0.043;
Matches 25; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
QY 1 GCNCAYGCNTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
DB 633 GCCCAGCGTTGGCCCAACCTTACAAAGACTATGAGTAAG 674
RESULT 43
ADL13501/c
ID ADL13501 standard; DNA; 167932 BP.

```
XX AC ADL13501;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #33.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX XX 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX XX WPI; 2003-559141/52.
XX DR
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PT a protein.
XX PS Disclosure; SEQ ID NO 33; 297bp; English.
XX CC
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 167932 BP; 42550 A; 42808 C; 42441 G; 40130 T; 0 U; 3 Other;
Query Match 64.5%; Score 21.4; DB 10; Length 167932;
Best Local Similarity 57.9%; Pred. No. 1.2e+02;
Matches 22; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
QY 1 GCNAYGCGTCCNTCCNTAYAAAGAYTAYGAGT 38
DB 130330 GCACACTCATTCTTCCTCTTAAATACTATGAGT 130293
RESULT 44
ADR84444
ID ADR84444 standard; DNA; 7880 BP.
XX AC ADR84444;
XX AC
XX DT 04-NOV-2004 (first entry)
XX DE Aspergillus fumigatus essential gene genomic sequence #255.
XX KW Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;
XX KW drug screening; ds.
XX KW
```

```
XX OS Aspergillus fumigatus.
XX PN WO2004067709-A2.
XX XX 12-AUG-2004.
XX PF 16-JAN-2004; 2004WO-US001099.
XX XX 17-JAN-2003; 2003US-0441281P.
XX PR 13-JUN-2003; 2003US-0478196P.
XX XX (ELIT-) ELITRA PHARM INC.
XX PA (ELIT-) ELITRA CANADA LTD.
XX PI Jiang B, Hu W, Lemieux S, Roemer T;
XX XX WPI; 2004-594200/57.
XX DR P-PSDB; ADR86205.
XX XX
XX PT New purified or isolated Aspergillus fumigatus nucleic acid molecule
XX PT encoding a gene product, useful for diagnosing and/or treating invasive
XX PT fungal infections, such as Farmer's lung disease.
XX XX
XX PS Claim 3; SEQ ID NO 255; 164pp; English.
XX CC
XX CC The present invention relates to Aspergillus fumigatus genes that are
XX CC essential and are potential targets for drug screening. The methods and
XX CC compositions of the present invention are useful for diagnosing and/or
XX CC treating invasive Aspergillus fumigatus infection, including the allergic
XX CC forms of the disease, such as Farmer's lung disease. They can also be
XX CC used in various drug discovery purposes, such as expression of the
XX CC recombinant protein, hybridization assay and construction of nucleic acid
XX CC arrays. The present sequence represents an Aspergillus fumigatus
XX CC essential gene full length genomic sequence, used during diagnosis and
XX CC drug development in the invention. These genes share a high degree of
XX CC sequence conservation with known essential genes of candida albicans. The
XX CC sequence data for this patent is not represented in the printed
XX CC specification, but was obtained in electronic format from WIPO.
XX SQ Sequence 7880 BP; 1905 A; 1902 C; 1920 G; 2153 T; 0 U; 0 Other;
Query Match 63.9%; Score 21.2; DB 13; Length 7880;
Best Local Similarity 57.1%; Pred. No. 62;
Matches 20; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 6 YCGTGGCCNTCCNTAYAAAGAYTAYGAGTNA 40
DB 1152 CCGTGGCCCTCTTCATATCAGGATCATGAGCTTA 1186
RESULT 45
ADQ56529
ID ADQ56529 standard; DNA; 421 BP.
XX AC ADQ56529;
XX XX
XX DT 21-OCT-2004 (first entry)
XX XX
XX DE Novel canine microarray-related DNA sequence SeqID7831.
XX KW canine microarray; drug screening; toxicity assay;
XX KW environmental pollutant; cellular response; gene expression profile;
XX KW toxic response; liver necrosis; fatty liver disease;
XX KW protein adduct formation; hepatitis; dog; ds.
XX OS Canis familiaris.
XX PN WO2004063324-A2.
XX PD 29-JUL-2004.
XX XX
XX PF 05-MAY-2003; 2003WO-US013853.
```

```
XX 03-MAY-2002; 2002US-0377240P.
XX (GENE-) GENE LOGIC INC.
XX (PFIZ ) PFIZER PROD INC.
XX
XX Diggins JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
XX
XX New isolated nucleic acid molecule, useful for drug screening and
XX toxicity assays or for assessing the impact, including toxicity, of a
XX compound, pharmaceutical agent or environmental pollutant on a cell or
XX living organism.
XX
XX Claim 1; SEQ ID NO 7831; 41pp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
XX sequences and the construction of canine microarrays containing a
XX significant portion of the canine genome. The isolated canine nucleic
XX acid sequences of the invention may be useful for drug screening and
XX toxicity assays. The invention is therefore useful for assessing the
XX impact, including toxicity, of a compound, pharmaceutical agent or
XX environmental pollutant on a cell or living organism. The methods are
XX useful for detecting genes that are up- or down-regulated in canines in a
XX disease state. The sequences are useful as diagnostic agents or markers
XX to detect a cellular response in a sample individually or as part of a
XX gene expression profile. It is also useful as a target for agents that
XX modulate gene expression or activity. The database is useful for
XX producing electronic Northern blots that allow the user to determine the cell
XX type or tissue in which a given gene is expressed and to allow
XX determination of the abundance or expression level of a given gene in a
XX particular tissue or cell. The methods are useful for determining the
XX similarity of a toxic response to one or more individual compounds. The
XX methods are useful for predicting at least one toxic response or the
XX likelihood that a compound or test agent will induce various specific
XX pathologies such as those of the liver (liver necrosis, fatty liver
XX disease, protein aduet or formation or hepatitis), those of the kidney,
XX heart, brain or testes, or other pathologies associated with at least one
XX of the toxins. The methods are also useful for predicting or elucidating
XX the potential cellular pathways influenced, induced or modulated by the
XX compound or test agent due to the similarity of the expression profile
XX compared to the profile induced by a known toxin. The present sequence is
XX that of a canine DNA sequence which was claimed for use during the
XX production of a canine microarray of the invention.
XX
XX Sequence 421 BP; 59 A; 147 C; 92 G; 87 T; 0 U; 36 Other;
XX
XX Query Match 62.7%; Score 20.8; DB 13; Length 421;
XX Best Local Similarity 55.0%; Pred. No. 41;
XX Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
XX
XX QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNA 40
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 301 GCCCATGCGAGGCCCTCACACCAACGAGCCGTACGCTCTTA 340
XX
XX RESULT 46
XX ABQ81845/C
XX ID ABQ81845 standard; DNA; 349980 BP.
XX
XX AC ABQ81845;
XX
XX XX 19-NOV-2002 (first entry)
XX
XX Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1101.
XX
XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
XX anti-diarrheic; antibacterial; inhibitor of Salmonella; detection;
XX identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
XX rotavirus; food composition; pharmaceutical composition; gene; ds.
XX
XX Bifidobacterium longum.
XX
```

```
OS Synthetic.
XX
XX EP1227152-A1.
XX
XX 31-JUL-2002.
XX
XX 30-JAN-2001; 2001EP-00102050.
XX
XX 30-JAN-2001; 2001EP-00102050.
XX
XX (NEST ) SOC PROD NESTLE SA.
XX
XX WPI; 2002-668397/72.
XX
XX Novel polynucleotide comprising Bifidobacterium genome sequence useful as
XX a probe or primer for detecting and/or identifying Bifidobacterium longum
XX in a biological sample.
XX
XX Disclosure; SEQ ID NO 1101; 80pp; English.
XX
XX The present invention describes a polynucleotide (I) comprising a
XX sequence of a Bifidobacterium genome selected from the nucleotide
XX sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
XX least 90% identity or which hybridises with the sequences given in
XX ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
XX fusion protein, comprising a sequence selected from 1097 sequences given
XX in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a
XX heterologous polypeptide. (I) has anti-diarrheic and antibacterial
XX activities, and can be used as an inhibitor of Salmonella. (I) (which is
XX a probe) is useful for the detection and/or identification of
XX Bifidobacterium longum in a biological sample. A carrier containing the
XX lactic acid bacterium Bifidobacterium longum NCC2705 (NCM I-2618) can be
XX used for preventing and/or treating diarrhoea brought about by pathogenic
XX bacteria and/or rotavirus. The carrier is a food composition selected
XX from milk, yogurt, curd, cheese, fermented milks, milk based fermented
XX products, ice-creams, fermented cereal based products, milk based
XX powders, infant formula, pet food or a pharmaceutical composition
XX selected from tablets, liquid bacterial suspensions, dried oral
XX supplement, wet oral supplement, dry tube feeding or wet tube feeding.
XX (I) is useful in DNA arrays or chips to carry out analysis of the
XX expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
XX Bifidobacterium related nucleotide sequences given in the Sequence
XX listing from the present invention but not mentioned further within the
XX specification. N.B. The sequence data for this patent is not represented
XX in the printed specification but is based on sequence information
XX supplied by the European Patent Office
XX
XX Sequence 349980 BP; 70780 A; 106600 C; 104724 G; 67876 T; 0 U; 0 Other;
XX
XX Query Match 62.7%; Score 20.8; DB 6; Length 349980;
XX Best Local Similarity 55.0%; Pred. No. 2.7e+02;
XX Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
XX
XX QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNA 40
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 129737 GCCCACCACCTGTCCCTCCAGTAGAAGGACTGCGAGCGGA 129698
XX
XX RESULT 47
XX AAT51538
XX ID AAT51538 standard; DNA; 2004 BP.
XX
XX AC AAT51538;
XX
XX XX 16-OCT-2003 (revised)
XX DT 23-APR-1997 (first entry)
XX
XX Mycobacterium gallisepticum incomplete pMGAL.5/1.6 gene.
XX
XX Adhesin; pMGAL.5; mycoplasma; diagnosis; vaccine; vector;
XX respiratory disease; poultry; haemagglutinin; promoter; probe; ds.
XX
XX Mycoplasma gallisepticum; strain S6.
XX
```

```
XX CA2135330-A.
PN
XX
XX 11-MAY-1995.
XX
XX 08-NOV-1994; 94CA-02135330.
XX
XX 10-NOV-1993; 93AU-00050593.
PR 20-APR-1994; 94US-00230312.
XX
XX (BROW/) BROWNING G F.
XX
XX Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;
XX WPI; 1995-241027/32.
DR
XX
XX New promoter region from a Mycoplasma gallisepticum adhesin gene - useful
PT when coupled to foreign antigen gene, for prodn. of multivalent live
PT vaccines, also new probes for detecting Mycoplasma and manipulating its
PT genome.
XX
XX Claim 2; Page 46-48; 81pp; English.
XX
XX A 10 kb DNA fragment was isolated from a Mycoplasma gallisepticum genomic
CC library using probes (AAT51533-34) based on tryptic peptides of a M.
CC gallisepticum PMGA adhesin. The fragment was divided into 5 putative
CC genes, PMGA1.2, PMGA1.3, PMGA1.4 and PMGA1.5/1.6 (AAT51535-38) on the
CC basis of deduced amino acids sequences. The 5 genes are closely related.
CC The regions between the genes (see also AAT51539-42) are highly
CC conserved. A consensus conserved sequence (AAT51532) can be used as a
CC probe to detect Mycoplasma infection, esp. in poultry. The promoter
CC sequences found in the conserved sequences may be linked to exogenous
CC antigen genes for prodn. of multivalent live vaccines. (Updated on 16-OCT
CC -2003 to standardise OS field)
XX
XX Sequence 2004 BP; 756 A; 317 C; 335 G; 596 T; 0 U; 0 Other;
SQ
Query Match 62.0%; Score 20.6; DB 2; Length 2004;
Best Local Similarity 52.6%; Pred. No. 79;
Matches 20; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
QY 5 AYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 818 ATGCTGGACAATCACCTAATAAGATAATGTAGTAAAG 855
RESULT 48
AAQ68672
ID AAQ68672 standard; DNA; 2577 BP.
XX
XX AAQ68672;
AC
XX
XX 28-FEB-1995 (first entry)
DT
XX
XX Partial PMGA1.5 and partial PMGA1.6 Mycoplasma gallisepticum DNA.
DE
XX
XX PMGA; adhesin gene complex; hemagglutinin; conserved sequences; primers;
KW probes; amplification; polymerase chain reaction; specific; detection;
KW PCR; 1.2; 1.3; 1.4; 1.5; 1.6; ss.
KW
XX
XX Mycoplasma gallisepticum.
OS
XX
XX Key Location/Qualifiers
FH
FT misc_feature 1. .1333
FT /tag= C
FT /note= "5'end of PMGA1.5"
FT misc_feature 61
FT /tag= a
FT /note= "unidentified residue"
FT misc_feature 1334. .1857
FT /tag= b
FT /note= "unidentified residues"
FT misc_feature 1857. .2557
```

```
FT
FT
XX
XX AU9350593-A.
XX
XX 26-MAY-1994.
XX
XX 10-NOV-1993; 93AU-00050593.
XX
XX 10-NOV-1992; 92AU-00005744.
XX
XX (UYME ) UNIV MELBOURNE.
XX
XX Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;
PI WPI; 1994-209061/26.
XX
XX Recombinant DNA constructs for Mycoplasma gallisepticum - for diagnosis,
PT treatment and prophylaxis of poultry respiratory disorders.
PT
XX Example 1; Fig 4; 51pp; English.
XX
XX AAQ68672 shows the 5'end of PMGA1.5 and the 3'end of PMGA1.6 DNA
CC sequences derived from a 10 kb insert from a Mycoplasma gallisepticum
CC genomic DNA library, detected by probes based on the T3 and C7 peptides.
CC Mycoplasma gallisepticum infection in poultry, humans and other animals
CC is of economic importance to many industries and it is desirable to
CC produce effective vaccines and probes for its detection. The sequences
CC and probes and vaccine vectors of the invention can be used for the
CC diagnosis and treatment of Mycoplasma gallisepticum infection, and for
CC prophylaxis
XX
XX Sequence 2577 BP; 772 A; 327 C; 340 G; 613 T; 0 U; 525 Other;
SQ
Query Match 62.0%; Score 20.6; DB 2; Length 2577;
Best Local Similarity 52.6%; Pred. No. 85;
Matches 20; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
QY 5 AYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 817 ATGCTGGACAATCACCTAATAAGATAATGTAGTAAAG 854
RESULT 49
AED71703
ID AED71703 standard; DNA; 1053 BP.
XX
XX AED71703;
AC
XX
XX 12-JAN-2006 (first entry)
DT
XX
XX Corynebacterium glutamicum MCP gene SEQ ID NO:1879.
DE
XX
XX marker and fine chemical production protein; screening; gene; ds.
KW
XX
XX Corynebacterium glutamicum.
OS
XX
XX US6962989-B1.
PN
XX
XX 08-NOV-2005.
PD
XX
XX 27-JUN-2000; 2000US-00605703.
PF
XX
XX 08-JUL-1999; 99US-0142764P.
PR
XX 03-SEP-1999; 99US-0152318P.
XX
XX (BADI ) BASF AG.
PA
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI WPI; 2005-743593/76.
XX
XX P-PSDB; AED71704.
DR
XX
```


CC enzymes. These are useful in the food, animal feed, cosmetics and
 CC pharmaceutical industries. The polynucleotides, optionally as primers and
 CC probes, can also be used for identification and classification of C.
 CC glutamicum and related species, e.g. for diagnosis, for genomic mapping,
 CC functional or evolutionary studies, gene manipulation and modulation of
 CC metabolic activity. Cells that containing the polynucleotides of the
 CC invention may produce fine chemicals in better yields, with higher
 CC productivity and/or more efficiently. NOTE: This sequence is not
 CC represented in the printed specification but is available in electronic
 CC format. The sequence represented in this record has been obtained from
 CC WO2003046123.

XX SQ Sequence 1201 BP; 275 A; 305 C; 331 G; 290 T; 0 U; 0 Other;

Query Match 61.4%; Score 20.4; DB 11; Length 1201;
 Best Local Similarity 52.5%; Pred. No. 85;
 Matches 21; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 GCNAYGCGTGGCCNTCCNTAYARGAYTAYGARGTNA 40
 |||||
 Db 197 GCGCAGCGCTTGACCGCCTTGTGAAGACGACCAAGTGA 236

RESULT 52
 ABQ76469
 ID ABQ76469 standard; cDNA; 2780 BP.

XX AC ABQ76469;
 XX DT 21-NOV-2002 (first entry)

XX DE S. cerevisiae BAX-associated cDNA fragment SEQ ID 363.

XX KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death; ss.

XX OS Saccharomyces cerevisiae.

XX PN WO200264766-A2.

XX PD 22-AUG-2002.

XX PF 21-DEC-2001; 2001WO-EP015398.

XX PR 22-DEC-2000; 2000EP-00870318.

XX PR 04-JAN-2001; 2001EP-00870002.

XX PR 09-JAN-2001; 2001EP-00870003.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;

XX DR WPI; 2002-667002/71.

XX DR P-PSDB; ABG93203.

XX PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.

XX PS Claim 36; Fig 1; 344pp; English.

XX CC This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cytostatic, fungicide; immunosuppressive, virucide and
 CC vasotropic activity and can be used in vaccines or for gene therapy. The
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,

CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenic flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polynucleotide associated with the Bax gene
 CC described in the disclosure of the invention

XX SQ Sequence 2780 BP; 969 A; 582 C; 479 G; 750 T; 0 U; 0 Other;

Query Match 61.4%; Score 20.4; DB 6; Length 2780;
 Best Local Similarity 56.8%; Pred. No. 1.1e+02;
 Matches 21; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 6 YGCVNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
 :|||||
 Db 2716 CGCCTGGCCATCACCATACATTGGTTAGGAGGAAAA 2752

RESULT 53
 AAH68525/c
 ID AAH68525 standard; DNA; 349980 BP.

XX AC AAH68525;

XX DT 26-SEP-2001 (first entry)

XX DE C glutamicum coding sequence fragment SEQ ID NO: 7060.

XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.

XX OS Corynebacterium glutamicum.

XX PN EF1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-00127688.

XX PR 16-DEC-1999; 99JP-00377484.

XX PR 07-APR-2000; 2000JP-00159162.

XX PR 03-AUG-2000; 2000JP-00280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX DR WPI; 2001-376931/40.

XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.

XX PS Disclosure; SEQ ID NO 7060; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain (II), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (III). (II) are also useful for identification of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents a L. pneumophila DNA.
XX
SQ Sequence 189430 BP; 56051 A; 34903 C; 40210 G; 58266 T; 0 U; 0 Other;
Query Match 60.8%; Score 20.2; DB 14; Length 189430;
Best Local Similarity 53.7%; Pred. No. 4.3e+02;
Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
QY 1 GCNCAYCGNTGCCNTCCNTAYARGAYTAYGAGTNA 41
DB 129266 GCCATAAGTTGCGCTGCTGCTATAAATAGGACGAGCAAA 129226
RESULT 60
AEB39171
ID AEB39171 standard; DNA; 263853 BP.
XX
AC AEB39171;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila DNA SEQ ID NO 3503.
XX
KW detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
PA (INSP) INST PASTEUR.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
XX
WPI; 2005-388305/40.
XX
XX
PT New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
PS Example 9; SEQ ID NO 3503; 660pp; English.
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (II), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (III). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or

CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents a L. pneumophila DNA.
XX
SQ Sequence 263853 BP; 83002 A; 52487 C; 46521 G; 81843 T; 0 U; 0 Other;
Query Match 60.8%; Score 20.2; DB 14; Length 263853;
Best Local Similarity 53.7%; Pred. No. 4.7e+02;
Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
QY 1 GCNCAYCGNTGCCNTCCNTAYARGAYTAYGAGTNA 41
DB 162843 GCCATAAGTTGCGCTGCTGCTATAAATAGGACGAGCAAA 162883
RESULT 61
ABL91353/c
ID ABL91353 standard; DNA; 669 BP.
XX
AC ABL91353;
XX
DT 29-AUG-2003 (revised)
DT 29-JUL-2002 (first entry)
XX
DE Chlamydia pneumoniae cp7041 ORF DNA, SEQ ID NO:340.
XX
KW Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
KW human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
KW strain CWL029; open reading frame; ORF; gene; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 1..669
FT /*tag= a
FT /product= "cp7041"
XX
PN WO200202606-A2.
XX
PD 10-JAN-2002.
XX
PF 03-JUL-2001; 2001WO-IB001445.
XX
PR 03-JUL-2000; 2000GB-00016363.
PR 11-JUL-2000; 2000GB-00017047.
PR 21-JUL-2000; 2000GB-00017983.
PR 07-AUG-2000; 2000GB-00019368.
PR 16-AUG-2000; 2000GB-00020440.
PR 14-SEP-2000; 2000GB-00022583.
PR 10-NOV-2000; 2000GB-00027549.
PR 22-DEC-2000; 2000GB-00031706.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Ratti G, Grandi G;
XX
WPI; 2002-154726/20.
DR N-PSDB; ABB90695.
XX
PT Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX
PS Claim 5; Page 176; 364pp; English.
XX
CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as

CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC DNA probe assay or blotting techniques for determining Chlamydia
CC pneumoniae gene expression. The present sequence represents a
CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 669 BP; 227 A; 103 C; 136 G; 203 T; 0 U; 0 Other;

Query Match 60.2%; Score 20; DB 6; Length 669;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 20; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 CNCAYCGTGGCCNTCCNTAYAAAGATYAGAR 36
DB 513 CTCATGCATAGGCTTTACCATAGAAAGATAACGAG 479

RESULT 62
AAS36060/c
ID AAS36060 standard; DNA; 5161 BP.
XX
AC AAS36060;
DT 17-DEC-2001 (first entry)
XX
DE Human cardiovascular system antigen genomic DNA SEQ ID No 1560.
XX
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antirheumatic; aniproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX
OS Homo sapiens.
XX
XX WO200153321-A2.
PN
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001340.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.


```
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226779P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0229437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-743766/70.
XX
XX New cardiovascular system related polynucleotides and polypeptides,
XX useful for preventing, treating, or ameliorating a medical condition,
XX such as cancer of cardiovascular tissues and cancer metastases.
XX
XX Claim 1; SEQ ID NO 1560; 262pp; English.
XX
XX The invention relates to human cardiovascular system related polypeptides
XX and the polynucleotides encoding them. The polypeptides, polynucleotides
XX and antibodies to the polypeptides are useful for diagnosing a
XX pathological condition or a susceptibility to a pathological condition,
XX for preventing, treating, or ameliorating a medical condition, such as
XX cancer of cardiovascular system tissues, proliferative disorders, foetal
XX and developmental abnormalities, haematopoietic disorders, diseases of
XX the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
XX arthritis), inflammation, allergies, neurological disorders (e.g.,
XX Alzheimer's disease, Parkinson's disease), cognitive disorders,
XX schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,
XX diabetes, atherosclerosis, cardiovascular disorders, angiogenic
XX disorders, kidney disorders, gastrointestinal disorders, pregnancy-
XX related disorders, endocrine disorders and infections. The nucleic acids
XX are also useful for chromosome identification, radiation hybrid mapping
XX or long-range restriction mapping. The polypeptides and polynucleotides
XX may also be used as food additives or preservatives to increase or
XX decrease storage capabilities, fat content or other nutritional
XX components. This sequence represents human cardiovascular system related
XX genomic DNA of the invention.
XX
XX Sequence 5161 BP; 1461 A; 1193 C; 1164 G; 1343 T; 0 U; 0 Other;
SQ
Query Match 60.2%; Score 20; DB 10; Length 5161;
Best Local Similarity 62.5%; Pred. No. 2e+02;
```

Matches	20;	Conservative	4;	Mismatches	8;	Indels	0;	Gaps	0;
Qy	4	CAYCGTGGCCNCCNCCNTAAYARGATYAGA	35						
Db	4936	CATGCTGGCGTCATATCTCTAATGATATGA	4905						
RESULT 64									
ADJ08172/C									
ID	ADJ08172	standard; DNA; 5161 BP.							
XX									
AC	ADJ08172;								
DT	04-NOV-2004	(first entry)							
XX									
DE		Human cardiovascular system associated polypeptide-related DNA SeqID1560.							
KW		autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;							
KW		breast neoplasms; liver neoplasm; cardiovascular disorder;							
KW		cardiac arrest; cerebrovascular disorder; cerebral ischaemia;							
KW		angiogenesis; nervous system disorder; Alzheimer's disease; infection;							
KW		ocular disorder; corneal infection; wound healing;							
KW		epithelial cell proliferation; skin aging; sunburn;							
KW		organ transplantation; cell culture; tissue regeneration; chemotaxis;							
KW		food additive; preservative; cardiovascular system associated antigen;							
KW		nuclear factor kappaB; NFkappaB; promoter element; human; ds.							
XX									
OS		Homo sapiens.							
PN		US2004005575-A1.							
XX									
PD		08-JAN-2004.							
XX									
PF		26-AUG-2002; 2002US-00227577.							
XX									
PR		31-JAN-2000; 2000US-0179065P.							
PR		04-FEB-2000; 2000US-0180628P.							
PR		24-FEB-2000; 2000US-0184663P.							
PR		02-MAR-2000; 2000US-0186350P.							
PR		16-MAR-2000; 2000US-0189874P.							
PR		17-MAR-2000; 2000US-0190076P.							
PR		18-APR-2000; 2000US-0198123P.							
PR		19-MAY-2000; 2000US-0205515P.							
PR		07-JUN-2000; 2000US-0209467P.							
PR		28-JUN-2000; 2000US-0211486P.							
PR		30-JUN-2000; 2000US-0215135P.							
PR		07-JUL-2000; 2000US-0216647P.							
PR		07-JUL-2000; 2000US-0216880P.							
PR		11-JUL-2000; 2000US-0217847P.							
PR		11-JUL-2000; 2000US-0217496P.							
PR		14-JUL-2000; 2000US-0218290P.							
PR		26-JUL-2000; 2000US-0220963P.							
PR		26-JUL-2000; 2000US-0220964P.							
PR		14-AUG-2000; 2000US-0224518P.							
PR		14-AUG-2000; 2000US-0224519P.							
PR		14-AUG-2000; 2000US-0225213P.							
PR		14-AUG-2000; 2000US-0225214P.							
PR		14-AUG-2000; 2000US-0225266P.							
PR		14-AUG-2000; 2000US-0225267P.							
PR		14-AUG-2000; 2000US-0225268P.							
PR		14-AUG-2000; 2000US-0225270P.							
PR		14-AUG-2000; 2000US-0225447P.							
PR		14-AUG-2000; 2000US-0225757P.							
PR		14-AUG-2000; 2000US-0225758P.							
PR		14-AUG-2000; 2000US-0225759P.							
PR		18-AUG-2000; 2000US-0226279P.							
PR		22-AUG-2000; 2000US-0226681P.							
PR		22-AUG-2000; 2000US-0226868P.							

```
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
PR 07-MAR-2002; 2002US-00091504.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2004-081713/08.
XX
XX New cardiovascular system-related nucleic acid molecule, useful for
XX diagnosing, preventing or treating diseases of the cardiovascular system,
XX and in chromosome mapping, drug screening or in pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 1560; 262pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
XX human cardiovascular system associated polypeptide (or antigens), or its
XX fragment. Also included recombinant vectors, recombinant host cells, an
XX isolated human cardiovascular system associated polypeptide (including
XX its fragment, allelic variant, species homologue or epitope), an isolated
XX antibody that binds specifically to a human cardiovascular system
XX associated polypeptide, diagnosing a pathological condition or
XX susceptibility to a pathological condition (comprising determining the
XX presence or absence of a mutation in human cardiovascular system
XX associated nucleic acid and diagnosing a condition based on the presence
XX or absence of the mutation), identifying a binding partner to human
XX cardiovascular system associated polypeptides, the gene corresponding to
XX the human cardiovascular system associated cDNA sequence and identifying
XX an activity in a biological assay comprising expressing the human
XX cardiovascular system associated cDNA in a cell, isolating the
XX supernatant, detecting an activity in a biological assay and identifying
XX the protein in the supernatant having the activity. The human
XX cardiovascular system associated nucleic acids and polypeptides are used
XX to prevent, treat or ameliorate a medical condition (for example in
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for
XX example autoimmune diseases such as rheumatoid arthritis,
XX hyperproliferative disorders, for example neoplasms of the breast or
XX liver, cardiovascular disorders, for example cardiac arrest,
XX
XX Query Match 60.2%; Score 20; DB 13; Length 5161;
XX Best Local Similarity 62.5%; Pred. No. 2e+02;
XX Matches 20; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
XX
XX 4 CAYGNTGGCNCNTCNCNTAYAAAGAYTAYGA 35
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db - 4936 CATGCTGGCGTCATATTCTAATGATTATGA 4905
```

```
RESULT 65
AAx91990_08
Continuation (9 of 13) of AAx91990 from base 800001 (Nucleotide sequence of the complete
WP Sequence split into 13 fragments LOCUS AAx91990 Accession AAx91990
WP Fragment Name Begin End
WP AAx91990_00 1 110000
WP AAx91990_01 100001 210000
WP AAx91990_02 200001 310000
WP AAx91990_03 300001 410000
WP AAx91990_04 400001 510000
WP AAx91990_05 500001 610000
WP AAx91990_06 600001 710000
WP AAx91990_07 700001 810000
WP AAx91990_08 800001 910000
WP AAx91990_09 900001 1010000
WP AAx91990_10 1000001 1110000
WP AAx91990_11 1100001 1210000
WP AAx91990_12 1200001 1230025

Query Match 60.2%; Score 20; DB 2; Length 110000;
Best Local Similarity 57.1%; Pred. No. 4.6e+02;
Matches 20; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 CNCAYGNTGGCNCNTCNCNTAYAAAGAYTAYGAR 36
|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 35280 CTCATGCATAGCGTTTACCATAGAAAGATAACGAG 35314

RESULT 66
ABD32968_6
Continuation (7 of 8) of ABD32968 from base 600001 (Human cancer-associated genomic DNA f
WP Sequence split into 8 fragments LOCUS ABD32968 Accession Abd32968
WP Fragment Name Begin End
WP ABD32968_0 1 110000
WP ABD32968_1 100001 210000
WP ABD32968_2 200001 310000
WP ABD32968_3 300001 410000
WP ABD32968_4 400001 510000
WP ABD32968_5 500001 610000
WP ABD32968_6 600001 710000
WP ABD32968_7 700001 779603

Query Match 60.2%; Score 20; DB 13; Length 110000;
Best Local Similarity 54.8%; Pred. No. 4.6e+02;
Matches 23; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCNCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10619 GGCCACTCATGCGCTTCACAAGAAAAGTGTAGGAAGAGAAA 10660

RESULT 67
ADY98830/C
ID ADY98830 standard; cDNA; 356 BP.
XX ADY98830;
XX
XX 16-JUN-2005 (first entry)
DT
DE T_reesei suppressive subtractive hybridization (SSH) cDNA clone Seq 184.
XX gene expression; suppressive subtractive hybridization; microarray;
XX microorganism; ss.
XX
XX Hypocrea jecorina.
XX
XX US2005069934-A1.
XX
XX 31-MAR-2005.
XX
XX 24-SEP-2004; 2004US-00950009.
XX
XX 25-SEP-2003; 2003US-0506140P.
XX
```


XX DE Plant full length insert polynucleotide seqid 33885.
XX AC
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX OS Unidentified.
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX PF 28-APR-2003; 2003US-00425114.
XX PR 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PA (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABA/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX DR
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX PS Claim 1; SEQ ID NO 33885; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX SQ Sequence 1083 BP; 267 A; 265 C; 321 G; 230 T; 0 U; 0 Other;
Query Match 59.0%; Score 19.6; DB 13; Length 1083;
Best Local Similarity 57.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 8 CNTGGCCNTCCNTAYARGAYTAYGARGTNA 40
DB 46 CGTGGCCCGCCCGTCTAAGGACTACGAGTTCA 78
RESULT 70
ADX36401

ID ADX36401 standard; cDNA; 1305 BP.
XX AC
XX ADX36401;
DT 21-APR-2005 (first entry)
XX DE Plant full length insert polynucleotide seqid 19221.
XX KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX OS Unidentified.
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX PF 28-APR-2003; 2003US-00425114.
XX PR 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PA (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABA/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX DR
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX PS Claim 1; SEQ ID NO 19221; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX SQ Sequence 1305 BP; 302 A; 350 C; 368 G; 285 T; 0 U; 0 Other;
Query Match 59.0%; Score 19.6; DB 13; Length 1305;
Best Local Similarity 57.6%; Pred. No. 2e+02;
Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 8 CNTGGCCNTCCNTAYARGAYTAYGARGTNA 40

Db 235 CGTCGCCGCCCGTCTAAGGACTACGAGTTCA 267

RESULT 71

ABZ90108

ID ABZ90108 standard; DNA; 286 BP.

XX AC ABZ90108;

XX DT 17-OCT-2003 (first entry)

XX DE Human nucleic acid sequence.

XX KW Human; antisense; lung dysfunction; nasal airway dysfunction; anti-inflammatory steroid; ubi-quinone; anti-inflammatory; antiallergic; antiasthmatic; hypotensive; immunosuppressive; cytosolic; gene therapy; antisense gene therapy; respiratory; lung; adenosine sensitivity; adenosine receptor; bronchodilation; bronchoconstriction; lung allergy; lung inflammation; respiratory disease; ds.

XX OS Homo sapiens.

XX PN WO200285308-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013135.

XX PR 24-APR-2001; 2001US-0286137P.

XX PA (EPIG-) EPIGENESIS PHARM INC.

XX PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;

XX PI Miller S, Tang L, Shahabuddin S;

XX DR WPI; 2003-229219/22.

XX PT Pharmaceutical composition for treating ailments associated with impaired respiration, has oligo(s) antisense to specific gene(s) or its corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or ubi-quinone.

XX PS Disclosure; SEQ ID NO 5350; 872pp; English.

XX CC The invention relates to a novel pharmaceutical composition, which has a first active agent comprising an oligonucleotide antisense to the initiation codon, coding region, 5' or 3' end genomic flanking regions, 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of junctions of genes encoding a polypeptide associated with lung and/or nasal airway dysfunction and a second active agent comprising an anti-inflammatory steroid and ubi-quinone. A composition of the invention has anti-inflammatory, antiallergic, antiasthmatic, hypotensive, immunosuppressive, and cytosolic activity. The composition may have a use in antisense gene therapy. The composition is useful for treating or preventing a respiratory, lung or malignant disease or condition, also for enhancing the prophylactic or therapeutic respiratory effect of an anti-inflammatory steroid in a subject, for reducing or depleting levels of, or reducing sensitivity to adenosine, reducing levels of adenosine receptor, producing bronchodilation, increasing levels of ubi-quinone or lung surfactant in a subject's tissue, or treating bronchoconstriction, lung inflammation, lung allergies, or a respiratory disease or condition.

XX CC Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 286 BP; 88 A; 65 C; 61 G; 72 T; 0 U; 0 Other;

Query Match 58.4%; Score 19.4; DB 10; Length 286;

Best Local Similarity 48.8%; Pred. No. 1.6e+02;

Matches 20; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 2 CNAYGCGTCCGTCNCCNTAARGAYTAYGARGTNAAR 42

Db 92 CACCTGCTGATATTCCTCACTACAAACATTTGGTGTCAA 132

RESULT 72

ABD26338

ID ABD26338 standard; DNA; 286 BP.

XX AC ABD26338;

XX DT 29-JUL-2004 (first entry)

XX DE AA459692 DNA fragment.

XX KW Human; antisense; bronchoconstriction; allergy; hyposecretion; pain; respiratory tract inflammation; adenosine sensitivity; lung; cancer; surfactant depletion; antiallergic; anti-inflammatory; antiasthmatic; analgesic; hypotensive; immunosuppressive; cytosolic; cystic fibrosis; beta-adrenergic agonist; respiratory disease; pulmonary vasoconstriction; respiratory distress syndrome; allergic rhinitis; pulmonary hypertension; emphysema; chronic obstructive pulmonary disease; cancer; bronchitis; pulmonary transplantation rejection; ds.

XX OS Homo sapiens.

XX PN WO200285309-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013143.

XX PR 24-APR-2001; 2001US-0286036P.

XX PA (EPIG-) EPIGENESIS PHARM INC.

XX PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;

XX PI Miller S, Tang L, Shahabuddin S;

XX DR WPI; 2003-093058/08.

XX PT Pharmaceutical composition for treating asthma, has antisense oligonucleotide containing less percentage of adenosine, targeted to nucleic acids associated with lung airway or lung dysfunction, and bronchodilating agent.

XX PS Claim 15; SEQ ID NO 5350; 763pp; English.

XX CC This invention describes a novel composition (a) a first active agent, comprising oligonucleotides, effective for alleviating bronchoconstriction, respiratory tract inflammation, allergies and reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors, surfactant depletion or hyposecretion, when administered to a mammal. The oligonucleotides are derived from a gene encoding or regulating expression of a target polypeptide associated with lung airway or lung dysfunction or cancer and can be anti-sense to the corresponding mRNA. The invention also describes a kit, that comprises: (a) a delivery device, in separate containers, (b) the oligonucleotides, (c) instructions for adding a carrier and for use of the kit. The composition of the invention has antiallergic, anti-inflammatory, antiasthmatic, analgesic, hypotensive, immunosuppressive and cytosolic activity, is a beta-adrenergic agonist. The composition is useful for preventing or treating a respiratory, lung or malignant disease. The administered composition comprises oligo and is administered to reduce the production or availability, or to increase the degradation of the target mRNA or to reduce the amount of target polypeptide present in the lungs. The pulmonary obstruction, and/or bronchoconstriction and/or lung inflammation, allergies and/or surfactant hypoproduction are associated with a disease or condition such as pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome, pain, cystic fibrosis, allergic rhinitis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary transplantation rejection, pulmonary infections, bronchitis or cancer. The reduced adenosine content of the anti-sense oligos corresponding to thymidines present in the target RNA serves to prevent the breakdown of

Query Match 58.4%; Score 19.4; DB 10; Length 614;
 Best Local Similarity 55.6%; Pred. No. 2e+02;
 Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAYCNCNTGCCNTCCNTAYAAARGAYTAYGAR 36
 DB 237 GCACACGATGGCCTTGCCAGACACAAACCAAGAG 272

RESULT 75
 ABZ57439/C
 ID ABZ57439 standard; cDNA; 918 BP.
 XX ABZ57439;
 XX AC
 XX AC
 XX AC
 DT 10-APR-2003 (first entry)
 XX
 DE Formamidopyrimidine-DNA glycosylase 11.77-encoding cDNA.
 XX
 KW Formamidopyrimidine-DNA glycosylase 11.77; recombinant production;
 KW gene therapy; malignant tumour; cancer; blood disease; HIV infection;
 KW human immunodeficiency virus; immune disorder; inflammatory condition;
 KW cytostatic; antiinflammatory; immunomodulator; gene; ss.
 XX
 OS Unidentified.
 XX
 XX
 FH Key Location/Qualifiers
 FH CDS 287..610
 FT FT /*tag= a
 FT FT /product= "Formamidopyrimidine-DNA glycosylase 11.77"
 FT FT
 XX
 XX CN1361246-A.
 XX
 XX
 XX 31-JUL-2002.
 XX
 XX 26-DEC-2000; 2000CN-00135918.
 XX
 XX 26-DEC-2000; 2000CN-00135918.
 XX
 XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 PA
 PA Mao Y, Xie Y;
 PI
 XX
 XX WPI; 2002-741868/81.
 DR
 DR P-PSDB; ABP58836.
 XX
 XX New polypeptide formamidopyrimidine-DNA glycosylase 11.77 and
 PT polynucleotides encoding this polypeptide.
 PT
 XX
 XX Claim 6; Page 25 (Disclosure); 32pp; Chinese.
 XX
 XX The invention relates to formamidopyrimidine-DNA glycosylase 11.77
 CC (ABP58836) and nucleic acids encoding it (ABZ57439). The protein has a
 CC molecular weight of 11.77 kD. The invention also relates to a method for
 CC the recombinant production of the protein, an antagonist of the protein,
 CC and the use of the protein, gene and antagonist in therapeutic
 CC applications. Formamidopyrimidine-DNA glycosylase 11.77 can be used in
 CC the treatment of a variety of diseases such as malignant tumours, blood
 CC diseases, HIV (human immunodeficiency virus) infection, immune disorders
 CC and inflammatory conditions. The present sequence represents cDNA
 CC encoding formamidopyrimidine-DNA glycosylase 11.77
 XX
 XX Sequence 918 BP; 164 A; 271 C; 271 G; 212 T; 0 U; 0 Other;

Query Match 58.4%; Score 19.4; DB 6; Length 918;
 Best Local Similarity 55.6%; Pred. No. 2.3e+02;
 Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAYCNCNTGCCNTCCNTAYAAARGAYTAYGAR 36
 DB 482 GCACACGATGGCCTTGCCAGACACAAACCAAGAG 447

```

XX OS Haemophilus influenzae.
XX PN WO2004078949-A2.
XX PD 16-SEP-2004.
XX PF 05-MAR-2004; 2004WO-US007001.
XX PR 06-MAR-2003; 2003US-0453134P.
XX PA (CHIL-) CHILDRENS HOSPITAL INC.
XX PI Bakaletz LO, Munson RS, Dyer DW;
XX WPI; 2004-662422/64.
XX PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
XX PT useful for treating or preventing NTHi bacterial infections of the middle
XX PT ear and/or nasopharynx.
XX PS Example 1; SEQ ID NO 11; 88pp; English.
XX CC The invention comprises nucleotide sequences (genes) from the genome of a
XX CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
XX CC sequences of the invention are useful for treating or preventing NTHi
XX CC bacterial infections of the middle ear and/or nasopharynx. The present
XX CC nucleic acid represents an NTHi contig sequence of the invention.
XX SQ Sequence 992 BP; 321 A; 173 C; 210 G; 287 T; 0 U; 1 Other;

Query Match 58.4%; Score 19.4; DB 13; Length 992;
Best Local Similarity 48.8%; Pred. No. 2.3e+02;
Matches 20; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 2 CCAAGCGTGGCCGTCCTCCCAATATGTTTACGAAGTAA 42
DB 939 CGCTTGATGCACATCCCGCAATATGTTTACGAAGTAA 899

RESULT 78
AAH99671
ID AAH99671 standard; cDNA; 1757 BP.
XX AC AAH99671;
XX DT 16-OCT-2001 (first entry)
XX DE Human protein encoding cDNA sequence SEQ ID NO:506.
XX KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX KW anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX KW antibacterial; endocrine; cardiant; central nervous system; virucide;
XX KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
XX KW antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
XX KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
XX KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX KW immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;
XX KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
XX KW allergic rhinitis; diabetes; multiple sclerosis; depression;
XX KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX KW neurological disorder; ss.
XX OS Homo sapiens.
XX PN WO200153455-A2.
XX PD 26-JUL-2001.
XX PF 22-DEC-2000; 2000WO-US035017.

```

```

XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457603/49.
XX P-PSDB; AAM25730.
XX PT Isolated human polynucleotides encoding polypeptides, useful for the
XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX PS Claim 1; Page 569; 1217pp; English.
XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX CC AAM25963. The proteins can have activities based on the tissues and cells
XX CC they are expressed in, such as: antiinflammatory; antirheumatic;
XX CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
XX CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
XX CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
XX CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
XX CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX CC encoding them can be used in gene therapy, antitense therapy and vaccine
XX CC production. The proteins and polynucleotides are useful for screening for
XX CC agonists or antagonists of a protein and for the treatment and diagnosis
XX CC of disorders associated with the activity of a protein e.g. inflammation,
XX CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX CC osteoporosis, severe combined immunodeficiency, eczema, allergic
XX CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
XX CC neurological disorders
XX SQ Sequence 1757 BP; 401 A; 526 C; 464 G; 366 T; 0 U; 0 Other;

Query Match 58.4%; Score 19.4; DB 4; Length 1757;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAVCNTGGCCNTCCNTATYAARGAYTAYGAR 36
DB 452 GCACACGCGATGGCTTGGCCAGACACAAACCAAGAG 487

RESULT 79
ABX71365/c
ID ABX71365 standard; cDNA; 2321 BP.
XX AC ABX71365;
XX DT 14-APR-2003 (first entry)
XX DE Human signal transduction-associated cDNA from clone DKFZphtes3_21d4.
XX KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX OS Homo sapiens.
XX PN WO200112659-A2.
XX PD 22-FEB-2001.
XX PF 18-AUG-2000; 2000WO-1B001496.
XX PR 18-AUG-1999; 99US-0149499P.
XX PR 28-SEP-1999; 99US-0156503P.
XX

```


PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX Wiemann S;
PI WPI; 2001-327840/34.
DR P-PSDB; ABUS3098.
XX
XX Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
XX Claim 1; Page 705; 1095pp; English.
XX
XX This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence encodes a polypeptide
CC described in the disclosure of the invention
XX
SQ Sequence 2321 BP; 461 A; 626 C; 739 G; 495 T; 0 U; 0 Other;
Query Match 58.4%; Score 19.4; DB 5; Length 2321;
Best Local Similarity 55.6%; Pred. No. 2.9e+02;
Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
QY 1 GCNCGCCTGCGCCNTCCNTAYAAAGAYTAYGAR 36
Db 1862 GCACGCGATGGCTTGGCCGACACAAACCAAGAG 1827
RESULT 80
ABZ11510/c
ID ABZ11510 standard; cDNA; 2416 BP.
XX
AC ABZ11510;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 392.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nontropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200270539-A2.
PN
XX 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2002-759812/82.
DR
XX P-PSDB; ABP69293.
XX
PT New polynucleotides comprising sequences assembled from expressed

PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
XX Claim 1; SEQ ID NO 392; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2416 BP; 457 A; 634 C; 748 G; 577 T; 0 U; 0 Other;
Query Match 58.4%; Score 19.4; DB 6; Length 2416;
Best Local Similarity 55.6%; Pred. No. 3e+02;
Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
QY 1 GCNCGCCTGCGCCNTCCNTAYAAAGAYTAYGAR 36
Db 1965 GCACGCGATGGCTTGGCCGACACAAACCAAGAG 1930
RESULT 81
ADM44028/c
ID ADM44028 standard; cDNA; 2416 BP.
XX
AC ADM44028;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #392.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN US2004053250-A1.
XX
PD 19-MAR-2004.
XX
XX 21-NOV-2002; 2002US-00302172.
XX
XX 05-MAR-2001; 2001US-00799451.
PR
XX 05-MAR-2002; 2002WO-US005095.
PR
XX 20-AUG-2002; 2002US-00225251.
XX
XX (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
XX
XX Tang YT, Xue A, Drmanac RT;
PI
XX WPI; 2004-238579/22.
DR
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
XX Disclosure; SEQ ID NO 392; 51pp; English.

XX			The invention relates to an isolated polynucleotide. The methods and compositions of the present invention are useful for the diagnosis and/or treatment of diseases or conditions associated with aberrant expression or activity of the arginine-rich protein-like polypeptides, such as cancer and inflammation. They can also be used in forensics, gene mapping, identification of mutations responsible for genetic disorders, and in assessing biodiversity. The present sequence represents a novel human arginine-rich protein cDNA.					
CC								
QQ	Sequence	2416 BP; 457 A; 634 C; 748 G; 577 T; 0 U; 0 Other;						
Query Match	58.4%; Score	19.4; DB	12; Length	2416;				
Best Local Similarity	55.6%; Pred.No.	3e+02;	Matches	20; Conservative	5; Mismatches	11; Indels	0; Gaps	0;
QY	1 GCNCAYCGTCGCCNTCNCNTAYAAAGGATYAYGAR	36						
DB	1965 GCACACCATGGCCCTTGCACGACACAACCAAGAG	1930						
	: : :							
RESULT 82								
AAC47167								
ID	AAC47167 standard; DNA; 3688 BP.							
XX								
AC	AAC47167;							
XX								
DT	18-OCT-2000 (first entry)							
XX								
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 52804.							
XX								
KW	Hybridisation assay; genetic mapping; gene expression control;							
KM	protein identification; signal transduction pathway; metabolic pathway;							
KW	promoter; termination sequence; ss.							
XX								
OS	Arabidopsis thaliana.							
XX								
PN	EP1033405-A2.							
XX								
PD	06-SEP-2000.							
XX								
Pf	25-FEB-2000; 2000EP-00301439.							
XX								
PR	25-FEB-1999; 99US-0121825P.							
PR	05-MAR-1999; 99US-0123180P.							
PR	09-MAR-1999; 99US-0123548P.							
PR	23-MAR-1999; 99US-0125788P.							
PR	25-MAR-1999; 99US-0126284P.							
PR	29-MAR-1999; 99US-0126785P.							
PR	01-APR-1999; 99US-0127462P.							
PR	06-APR-1999; 99US-0128234P.							
PR	08-APR-1999; 99US-0128714P.							
PR	16-APR-1999; 99US-0129845P.							
PR	19-APR-1999; 99US-0130077P.							
PR	21-APR-1999; 99US-0130449P.							
PR	23-APR-1999; 99US-0130510P.							
PR	23-APR-1999; 99US-0130891P.							
PR	28-APR-1999; 99US-0131449P.							
PR	30-APR-1999; 99US-0132048P.							
PR	30-APR-1999; 99US-0132407P.							
PR	04-MAY-1999; 99US-0132484P.							
PR	05-MAY-1999; 99US-0132485P.							
PR	06-MAY-1999; 99US-0132486P.							
PR	06-MAY-1999; 99US-0132487P.							
PR	07-MAY-1999; 99US-0132863P.							
PR	11-MAY-1999; 99US-0134256P.							
PR	14-MAY-1999; 99US-0134218P.							
PR	14-MAY-1999; 99US-0134219P.							
PR	14-MAY-1999; 99US-0134221P.							
PR	14-MAY-1999; 99US-0134370P.							
PR	18-MAY-1999; 99US-0134376P.							
PR	19-MAY-1999; 99US-0134941P.							
PR	20-MAY-1999; 99US-0135124P.							

```
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 58.4%; Score 19.4; DB 3; Length 3688;
Best Local Similarity 55.6%; Pred. No. 3.4e+02;
Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 7 GCNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
    ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1240 GCTGGACCAACCGTATACAGACTACGATGCAAG 1275

RESULT 83
ADFO1135
ID ADF01135 standard; DNA; 3738 BP.
XX
AC ADF01135;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacterial polynucleotide #1420.
XX
KW Proteus mirabilis infection; bacterial infection; antibacterial;
immunostimulant; gene; db.
XX
OS Proteus mirabilis.
XX
PN US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
(GENO-) GENOME THERAPEUTICS CORP.
PA
PI Breton GL;
XX
WPI: 2003-895291/82.
P-PSDB; ADF05307.
XX
New Proteus mirabilis polypeptides and polynucleotides, useful as
reagents for diagnosis of bacterial disease, as components of
antibacterial vaccines, as targets for antibacterial drugs, or as
biocontrol agents for plants.
XX
Disclosure; SEQ ID NO 1420; 870pp; English.
XX
The invention relates to new Proteus mirabilis polypeptides and
polynucleotides. The invention also relates to antibodies against the
polypeptides, methods for producing the polypeptides, a method of
generating vaccines for immunising an individual against P. mirabilis, a
method for evaluating a compound for the ability to bind a P. mirabilis
polypeptide and a method for screening test compounds for anti-bacterial
activity. The polypeptides and polynucleotides are useful as molecular
targets for diagnosing, preventing and treating pathological conditions
resulting from bacterial infection, as reagents for diagnosis of
bacterial diseases, as components of antibacterial vaccines, as targets
for antibacterial drugs or as bio-control agents for plants. This
sequence represents a Proteus mirabilis polynucleotide of the invention.
XX
Sequence 3738 BP; 1241 A; 730 C; 722 G; 1045 T; 0 U; 0 Other;

Query Match 58.4%; Score 19.4; DB 10; Length 3738;
Best Local Similarity 52.6%; Pred. No. 3.4e+02;
Matches 20; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 5 AYGCTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
    ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3278 AGCATGGATTGCACACTTCAAGATTACAAAGTGAAA 3315
```

RESULT 84

AZ1190/c
ID AZ1190 standard; cDNA; 418 BP.

XX XX
AC AC
AAZ1190;
XX DT
DT 04-NOV-1999 (first entry)
XX DE
DE Interleukin-3 coding sequence from b2HFLS20W cDNA library.
XX KW
KW Interleukin-3; IL-3; human; bone marrow; blood cell; immunodeficiency;
XX infection; immune disorder; rheumatoid arthritis; chemotherapy; therapy;
KW myeloid cell activation; lymphoid cell; leukopaenia; dendritic-like cell;
KW haematopoietic precursor cell; vaccine development; ss.
XX OS
OS Homo sapiens.

XX FH
FH Key Location/Qualifiers
FT CDS 2..202
FT FT /tag= a
FT PT /product= "Interleukin-3"
XX PN
PN WO9941382-A2.
XX PD
PD 19-AUG-1999.
XX PF
PF 17-FEB-1999; 99WO-US001484.
XX PR
PR 17-FEB-1998; 98US-00024820.
XX PR
PR 23-OCT-1998; 98US-00177467.
XX PR
PR 11-DEC-1998; 98US-00209534.
XX PA
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Crkvenjakov R, Dickson M, Drmanac S, Labat I,
PI Leshkowitz D, Kika D, Ford J;
XX DR
DR WPI; 1999-527369/44.
XX P-PSDB; AAV32936.
XX PT
PT New interleukin-3 useful for treating conditions requiring increases in
PT numbers or differentiation of hematopoietic cells, e.g. autoimmune
XX diseases, leucopenia.
XX PS
PS Claim 1; Fig 2; 158pp; English.

XX CC
CC This sequence encodes the novel interleukin-3 (IL-3) of the invention.
XX CC
CC This sequence was isolated from the human foetal liver-spleen tissue cDNA
XX library b2HFLS20W. The IL-3 polypeptides are useful in vitro for
XX stimulating growth and activation of bone marrow and blood cells before
XX infusion of these cells into patients. They can be used to treat various
XX immunodeficiencies (e.g. resulting from viral infection) or immune
XX disorders (e.g. rheumatoid arthritis) or other conditions in which
XX haematopoietic cell populations have been reduced/destroyed e.g. due to
XX radiation or chemotherapy. They are especially useful to activate mature
XX myeloid and/or lymphoid cells and especially to treat leukopaemia. The
XX polypeptides can also be used to stimulate differentiation in
XX haematopoietic precursor cells, especially haematopoietic stem or
XX progenitor cells, lymphoid cells or myeloid (especially monocyte) cells,
XX e.g. the differentiation of haematopoietic precursor cells into dendritic-
XX like cells. This is useful in vivo to increase the number of dendritic-
XX e.g. to treat pathogenic infections, autoimmune diseases and in
XX transplantation; in vitro production of dendritic cells is useful
XX clinically (e.g. in vaccine development) since isolation of these cells
XX from tissues is difficult. IL-3 can be combined with a carrier in
XX compositions useful as above, and used to generate antibodies useful in
XX assays and therapeutically. The polynucleotides are useful e.g. as
XX hybridisation probes, as oligomers for PCR, for chromosome and gene
XX mapping and for polypeptide production

Query Match 57.8%; Score 19.2; DB 2; Length 418;
Best Local Similarity 53.8%; Pred. No. 2.2e+02;
Matches 21; Conservative 5; Mismatches 13; Indels 0; Gaps 0

QY 4 CAYGCGTGGCNCNTCCNTAAGATTAAGGATTAGGAATAAAG 42
DB 82 CATGCTGGGTCTTCCTTAAGAGGATTAGGAATAAAG 44

RESULT 85

AAD39735/c
ID AAD39735 standard; DNA; 418 BP.

XX AC
AC AAD39735;
XX DT
DT 22-OCT-2002 (first entry)
XX DE
DE Human interleukin-3 (IL-3) DNA #1.
XX KW
KW Human; Interleukin-3; IL-3; gene mapping; chromosome mapping; radiation;
XX haematopoiesis; chemotherapy; pathogen infection; autoimmune disease;
KW transplant rejection; aplastic anaemia; leukopaenia; immunosuppressive;
KW leukaemia; cytostatic; immunostimulant; haemostatic; ds.
XX OS
OS Homo sapiens.
XX US
US US2002058018-A1.
XX PD
PD 16-MAY-2002.
XX PF
PF 23-FEB-2001; 2001US-00792246.
XX PR
PR 17-AUG-1999; 99US-00376732.
XX PA
PA (HYSE-) HYSEQ INC.
XX PI
PI Ford J;
XX DR
DR WPI; 2002-499505/53.
XX PT
PT New nucleic acid encoding a human interleukin-3, useful for promoting
XX growth and differentiation of hematopoietic precursor cells, e.g. for
XX treating leucopenia.
XX PS
PS Claim 1; Fig 2; 58pp; English.

XX CC
CC The present invention relates to human interleukin-3 (IL-3) proteins and
XX polynucleotides encoding such proteins. Sequences of the invention are
XX useful for gene mapping, chromosome mapping and to produce antisense
XX sequences. They are used to promote survival, growth and differentiation
XX of haematopoietic precursor cells, for the treatment of any condition
XX associated with inadequate haematopoiesis, e.g. side effects of radiation
XX or chemotherapy, pathogen infections, autoimmune diseases, transplant
XX rejection, aplastic anaemia, leukaemia, leukopaenia and in screening for
XX specific agonists or antagonists, potential therapeutic agents. The
XX present sequence is human IL-3 DNA

SQ Sequence 418 BP; 97 A; 111 C; 109 G; 101 T; 0 U; 0 Other;

Query Match 57.8%; Score 19.2; DB 6; Length 418;
Best Local Similarity 53.8%; Pred. No. 2.2e+02;
Matches 21; Conservative 5; Mismatches 13; Indels 0; Gaps 0

QY 4 CAYGCGTGGCNCNTCCNTAAGATTAAGGATTAGGAATAAAG 42
DB 82 CATGCTGGGTCTTCCTTAAGAGGATTAGGAATAAAG 44

RESULT 86

AAD39736/c
ID AAD39736 standard; DNA; 548 BP.

Qy 4 CAYGCNTGGCCNTCNCCNTAYAAARGAYTAYGARGTNAAR 42
|||:||||| ||| |||:||||| |||: |||:
Db 81 CATGCCTGGGTCTTCCCTAAGAGGATTAGGAAATAAAG 43

RESULT 88
AAC57760/c
ID AAC57760 standard; DNA: 1001 BP.

RESULT 89	
AD030325	
ID	AD030325 standard; CDNA; 1068 BP.
XX	
AC	AD030325;
XX	
DT	29-JUL-2004 (first entry)
XX	
DE	Mouse GPCR RAI3 polynucleotide, SEQ ID NO:1428.
XX	
KW	G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW	transgenic mouse; neurological disorder; adrenal gland disorder;
KW	colon disorder; intestinal disorder; cardiovascular disorder;
KW	muscular disorder; blood disorder; immune disorder; bone disorder;
KW	joint disorder; metabolic disorder; nutritive disorder; cancer;
KW	kidney disorder; liver disorder; lung disorder; breast disorder;
KW	ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW	skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW	thyms disorder; thyroid disorder; antiparkinsonian; antineoplastic;
KW	cytostatic; antinflammatory; vasotropic; antianginal; antiarrhythmic;
KW	CNS; central nervous system; respiratory; antiarrhythmic; antidiabetic;
KW	viricide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW	dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW	immunosuppressive; nephrotropic; Gene therapy; GPCR modulator; mouse;
KW	murine; Gene; ss.
OS	
XX	Mus musculus.
XX	
PN	W02004040000-A2.
XX	
PD	13-MAY-2004.
XX	
PF	09-SEP-2003; 2003WO-US028226.
XX	
PR	09-SEP-2002; 2002US-0409303P.
PR	09-APR-2003; 2003US-0461329P.
XX	
PA	(PRIM-) PRIMAL INC.
XX	
PI	Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI	Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
XX	
XX	WFI; 2004-390329/36.
DR	P-PSDB; ADO29633.
DR	
XX	
PT	Novel mammalian G protein coupled receptors, useful for identifying
PT	compounds that modulates diagnosing and treating disease condition
PT	associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT	pectoris, Parkinson's disease.
XX	
PS	Claim 151; SEQ ID NO 1428; 542pp; English.
XX	
CC	The invention relates to human and mouse G protein-coupled receptors
CC	(GPCRs) and nucleic acids encoding them. The invention also relates to
CC	sequences at least 90% identical to the GPCR proteins and nucleic acids
CC	of the invention; methods of treating, preventing or diagnosing diseases
CC	associated with GPCRs of the invention; methods of screening for
CC	compounds useful in the treatment of GPCR-related diseases; a transgenic
CC	mouse comprising a GPCR gene of the invention; a mouse comprising a
CC	mutation in a GPCR transgene or in an endogenous GPCR gene; cells deriv-
CC	ed from the transgenic mice; kits comprising several mice, each of which has
CC	a mutation in a different GPCR gene of the invention; and kits comprising
CC	probes which hybridise to GPCR polynucleotides of the invention. The
CC	invention further discloses variants of the GPCR polypeptides and vectors
CC	comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC	be used in the diagnosis, treatment or prevention of a wide variety of
CC	diseases including neurological disorders (e.g., Alzheimer's disease,
CC	depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC	disorders of the adrenal gland; disorders of the colon or intestine
CC	(e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC	syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC	myocardial infarction); muscular disorders; blood disorders (e.g.,
CC	anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or


```
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465558/50.
XX p-PSDB; AAU21715.
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid arthritis.
XX
XX Claim 4; SEQ ID NO 158; 687pp; English.
XX
XX The present invention relates to the isolation of novel human neoplastic
CC
```

```
CC disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC sequences encoding for these polypeptides. The sequences of the invention
CC are useful in the diagnosis, treatment, prevention and/or prognosis of
CC disorders involving neoplastic disease such as hyperproliferative
CC glioma, adult liver cancer, childhood cerebellar astrocytoma, or
CC Hodgkin's lymphoma). The sequences of the invention may also be useful
CC for treating other disorders such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, pulmonary disorders, cardiovascular disorders and renal
CC disorders. The polynucleotide sequences of the invention are also useful
CC in gene therapy. AAS34767-AAS35050 represent cDNA sequences encoding for
CC the novel human neoplastic disease associated polypeptides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 2648 BP; 903 A; 384 C; 499 G; 860 T; 0 U; 2 Other;
Query Match 57.8%; Score 19.2; DB 4; Length 2648;
Best Local Similarity 56.8%; Pred. No. 3.8e+02;
Matches 21; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
QY 5 AYGCTGCGCTCCNCCNTAYARGAYTAYGARGTNA 41
Db 1245 ACGCTGGACTTCTCAAAATATGAGCATGAGTCAA 1209
RESULT 96
ADC46072/c
ID ADC46072 standard; cDNA; 2648 BP.
AC ADC46072;
XX
XX 18-DEC-2003 (first entry)
DE Human neoplastic disease-associated gene 148 cDNA #1.
XX
XX Neoplastic disease-associated polypeptide; gene therapy; diabetes;
XX hyperproliferative disease; cancer; autoimmune disorder; multiple sclerosis;
XX rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
XX autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
XX haematologic disorder; anaemia; thrombocytopaenia; allergic reaction;
XX asthma; eczema; inflammatory disorder; ischaemia-reperfusion injury;
XX inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
XX Alzheimer's disease; Parkinson's disease; renal disorder;
XX acute glomerulonephritis; end-stage renal disease;
XX cardiovascular disorder; atherosclerosis; myocarditis;
XX infectious disease; AIDS; cachexia; anorexia; wound healing;
XX epithelial cell proliferation; Human; ss.
XX
XX Homo sapiens.
XX
XX US2003082758-A1.
XX
XX 01-MAY-2003.
XX
XX 22-MAR-2002; 2002US-00103313.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184564P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
```


CC condition or a susceptibility to a pathological condition in a subject,
 CC identifying a binding partner to the polypeptide, identifying an activity
 CC in a biological assay, and the gene corresponding to the cDNA sequence.
 CC The polypeptides, polynucleotides and antibodies are useful for
 CC detecting, preventing, diagnosing, prognosticating, treating or
 CC ameliorating medical conditions such as hyperproliferative diseases or
 CC cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis,
 CC systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis
 CC or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.
 CC anaemia or thrombocytopaenia), allergic reactions including asthma or
 CC eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,

Query Match 57.8%; Score 19.2; DB 10; Length 2648;
 Best Local Similarity 56.8%; Pred. No. 3.8e+02;
 Matches 21; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
 Qy 5 AYGNTGGCNCNTCCNTAAAGAYTAYGARGTNA 41
 Db 1245 ACGCCTGGACTTCTCAAAATTTATGAGCATGAAGTCAA 1209

RESULT 97
 ACA30557/c

ID ACA30557 standard; DNA; 3096 BP.

AC ACA30557;

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #12214.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX drug design; gene.

OS Campylobacter jejuni.

PN W0200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 23-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362659P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU26687.

XX Claim 14; SEQ ID NO 18427; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 3096 BP; 1189 A; 420 C; 642 G; 845 T; 0 U; 0 Other;

Query Match 57.8%; Score 19.2; DB 8; Length 3096;

Best Local Similarity 52.5%; Pred. No. 4e+02;

Matches 21; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GCNCAYGCNTGCCNTCCNTCCNTAAAGAYTAYGARGTNA 40

Db 1978 GCCCTTCATAGGCTTATCTATATATAAGAGTATGAGCGA 1939

RESULT 98

ADV35091

ID ADV35091 standard; cDNA; 3397 BP.

AC ADV35091;

DT 10-FEB-2005 (first entry)

XX Rat cDNA differentially expressed in the presence of VIP SeqID167.

DE rat; VIP; vasoactive intestinal polypeptide; ss;

XX multi-paramater high throughput screening; MPHTS; disease signature;

XX neuropsychiatric; neurodegenerative; schizophrenia;

XX bipolar affective disorder; BAD; autism; Parkinson's;

XX Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.

XX *Rattus norvegicus*.

XX US2003096264-A1.

XX 22-MAY-2003.

XX 18-JUN-2002; 2002US-00175523.

XX 18-JUN-2001; 2001US-0299151P.

PR 07-SEP-2001; 2001US-0317828P.

PR 25-SEP-2001; 2001US-0325150P.

PR 14-NOV-2001; 2001US-0333047P.

PR 18-JAN-2002; 2002US-034936P.

PR 04-MAR-2002; 2002US-0361834P.

XX (PSYC-) PSYCHIATRIC GENOMICS INC.

XX Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;

PI Palfreyman M, Rajan P;

XX WPI; 2004-118903/12.

XX Identifying a compound that can treat disease or disorders, such as, a
 PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
 PT determining the expression of one or more efficacy genes in a cell

PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
PA (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX WPI; 2003-148464/14.
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
SQ Sequence 3477 BP; 720 A; 1050 C; 998 G; 709 T; 0 U; 0 Other;
Query Match 57.8%; Score 19.2; DB 10; Length 3477;
Best Local Similarity 52.5%; Pred. No. 4.1e+02;
Matches 21; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
OY 2 CMCAYGCTGGCCNTCCNCTAARGAYTAYGARGTNA 41
DB 2510 CGGATGAAGGGTCTCTCCCTCTACCACGCTCTATGAGGTGAA 2549

Search completed: June 10, 2006, 15:33:27
Job time : 348.45 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:19:28 ; Search time 5182.8 seconds
(without alignments)
453.155 Million cell updates/sec

Title: US-10-600-816-33

Perfect score: 33.2

Sequence: 1 gcncaaycgtggcncntcnc.....ayaargaytaygargtnaar 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

EST:**

1: gb_est1:**

2: gb_est3:**

3: gb_est4:**

4: gb_est5:**

5: gb_est6:**

6: gb_est7:**

7: gb_est8:**

8: gb_est9:**

9: gb_est10:**

10: gb_est11:**

11: gb_gss1:**

12: gb_gss2:**

13: gb_gss3:**

14: gb_gss4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	90.4	158	2	BI063311 IL3-UT011
C 2	30	90.4	307	7	BE818833 CM2-BN030
C 3	30	90.4	341	7	BE818817 CM2-BN030
C 4	30	90.4	384	7	BE818835 CM2-BN030
C 5	30	90.4	396	7	BE818828 CM2-BN030
C 6	30	90.4	404	7	BE818840 CM2-BN030
C 7	30	90.4	424	7	BE818805 CM2-BN030
C 8	30	90.4	449	7	BE818807 CM2-BN030
C 9	30	90.4	453	5	CD672930 f18a05.Y
C 10	30	90.4	458	7	BE818871 CM2-BN030
C 11	30	90.4	466	7	BE818859 CM2-BN030
C 12	30	90.4	467	7	BE818852 CM2-BN030
C 13	30	90.4	571	3	BP253378 BP253378
C 14	30	90.4	577	4	CB152636 K-EST0209
C 15	30	90.4	580	3	BP252234 BP252234
C 16	30	90.4	580	3	BP267627 BP267627
C 17	30	90.4	580	3	BP268151 BP268151
C 18	30	90.4	582	3	BP256144 BP256144
C 19	30	90.4	582	3	BP363758 BP363758

BP256147	BP256147	584	30	90.4	3	BP256147
CB130950	CB130950	621	30	90.4	7	CB130950
BE747107	BE747107	630	30	90.4	4	BE747107
BM843051	BM843051	637	30	90.4	3	BM843051
BF586108	BF586108	699	30	90.4	7	BF586108
BG251131	BG251131	702	30	90.4	2	BG251131
BQ887352	BQ887352	851	30	90.4	3	BQ887352
CA454746	CA454746	865	30	90.4	4	CA454746
BM011054	BM011054	877	30	90.4	2	BM011054
BQ958341	BQ958341	899	30	90.4	3	BQ958341
EG831564	EG831564	901	30	90.4	2	EG831564
BX372160	BX372160	918	30	90.4	4	BX372160
BU543952	BU543952	923	30	90.4	3	BU543952
BX370558	BX370558	962	30	90.4	4	BX370558
BQ67433	BQ67433	1023	30	90.4	6	BQ67433
CR597125	CR597125	1830	30	90.4	6	CR597125
BI193620	BI193620	1125	30	90.4	2	BI193620
CF780868	CF780868	916	30	90.4	5	CF780868
BE818831	BE818831	365	30	90.4	7	BE818831
BQ685729	BQ685729	940	30	90.4	3	BQ685729
AU140676	AU140676	680	30	90.4	1	AU140676
AA615518	AA615518	454	30	90.4	1	AA615518
CA762213	CA762213	652	30	90.4	4	CA762213
AL210083	AL210083	952	30	90.4	14	AL210083
AA222951	AA222951	336	30	90.4	1	AA222951
W10531	W10531	493	30	90.4	10	W10531
AL214950	AL214950	944	30	90.4	14	AL214950
CD517976	CD517976	1212	30	90.4	5	CD517976
DT989670	DT989670	1439	30	90.4	10	DT989670
BE006062	BE006062	583	30	90.4	7	BE006062
QA692009	QA692009	560	30	90.4	11	QA692009
CL815091	CL815091	805	30	90.4	13	CL815091
AG837338	AG837338	842	30	90.4	14	AG837338
CV968179	CV968179	880	30	90.4	8	CV968179
CD190259	CD190259	315	30	90.4	5	CD190259
DU888204	DU888204	936	30	90.4	14	DU888204
CE724340	CE724340	626	30	90.4	12	CE724340
DX143003	DX143003	661	30	90.4	14	DX143003
DX134811	DX134811	734	30	90.4	14	DX134811
CZ983576	CZ983576	859	30	90.4	13	CZ983576
DR753481	DR753481	978	30	90.4	10	DR753481
CK084463	CK084463	583	30	90.4	5	CK084463
BZ990137	BZ990137	756	30	90.4	12	BZ990137
CC716597	CC716597	774	30	90.4	12	CC716597
CB673669	CB673669	792	30	90.4	4	CB673669
CB684312	CB684312	820	30	90.4	4	CB684312
CB649010	CB649010	821	30	90.4	4	CB649010
CB668485	CB668485	842	30	90.4	4	CB668485
BY321580	BY321580	343	30	90.4	4	BY321580
BI861966	BI861966	615	30	90.4	2	BI861966
CF585003	CF585003	1043	30	90.4	5	CF585003
AQ461894	AQ461894	544	30	90.4	11	AQ461894
DX336552	DX336552	573	30	90.4	14	DX336552
AQ727367	AQ727367	583	30	90.4	11	AQ727367
CW064489	CW064489	602	30	90.4	13	CW064489
CN249531	CN249531	615	30	90.4	8	CN249531
CZ770405	CZ770405	659	30	90.4	13	CZ770405
DU806777	DU806777	686	30	90.4	13	DU806777
CW154436	CW154436	687	30	90.4	14	CW154436
DX201680	DX201680	751	30	90.4	14	DX201680
DM547742	DM547742	794	30	90.4	10	DM547742
CW825776	CW825776	813	30	90.4	13	CW825776
CZ881566	CZ881566	843	30	90.4	13	CZ881566
BE548141	BE548141	912	30	90.4	7	BE548141
BP701218	BP701218	369	30	90.4	3	BP701218
AV802339	AV802339	435	30	90.4	7	AV802339
AN177524	AN177524	498	30	90.4	11	AN177524
DN752781	DN752781	581	30	90.4	9	DN752781
AL740353	AL740353	590	30	90.4	14	AL740353
CD496729	CD496729	647	30	90.4	5	CD496729
CW098794	CW098794	669	30	90.4	13	CW098794
CW309783	CW309783	677	30	90.4	13	CW309783
CW309782	CW309782	685	30	90.4	13	CW309782

93	20.8	62.7	861	4	CA790172	AGENCOURT	C 166	20.2	60.8	227	4	CA522568	KA512007H0
94	20.8	62.7	906	10	DW670971	CNB362-G1	167	20.2	60.8	256	1	AA650481	ns98f09.s
95	20.8	62.7	967	4	CB209690	AGENCOURT	C 168	20.2	60.8	386	11	AQ076024	CIT-HSP-2
96	20.8	62.7	974	8	CV922581	Phrpch 20	C 169	20.2	60.8	391	1	AU302351	AU302351
97	20.8	62.7	1047	5	CD242359	AGENCOURT	C 170	20.2	60.8	403	3	BP489700	BP489700
98	20.8	62.7	1166	5	CD497911	CD430-E11	C 171	20.2	60.8	442	4	CA525074	CA525074
99	20.8	62.7	1175	5	CD499520	CD430-G08	C 172	20.2	60.8	459	4	CA523402	CA523402
100	20.8	62.7	1226	9	DW670301	CFW55-G12	C 173	20.2	60.8	459	10	H01408	Y199b10.s1
101	20.8	62.7	1346	10	DT950356	CFW109-A0	C 174	20.2	60.8	539	11	AQ253989	HS 2196.B
102	20.8	62.7	1379	9	DN684833	CGX38-C06	C 175	20.2	60.8	540	10	DV217202	D1-K1992
103	20.8	62.7	1405	10	DT993768	CNB169-A0	C 176	20.2	60.8	612	11	AQ562093	HS 5233.A
104	20.8	62.7	1405	14	AG310602	Mus_muscu	C 177	20.2	60.8	665	4	CA012966	HT06P18r_
105	20.6	62.0	269	10	DW431455	HHA0E0301	C 178	20.2	60.8	667	2	BF949531	MR3-NN021
106	20.6	62.0	422	13	DUI39603	OG_Aba002	C 179	20.2	60.8	674	12	CC313716	TAM32-121
107	20.6	62.0	509	14	CR241190	Forward s	C 180	20.2	60.8	775	14	CT411387	Sus scrof
108	20.6	62.0	585	14	DX110143	OG_Aba013	C 181	20.2	60.8	259	13	CL416588	ZMMBB043
109	20.6	62.0	620	13	CZ779324	OC_Ba014	C 182	20.2	60.8	310	7	BB504145	BB504145
110	20.6	62.0	627	14	DU660987	OG_Aba006	C 183	20.2	60.8	350	12	CR027938	tigr-gss-
111	20.6	62.0	635	14	DU996725	OG_Aba011	C 184	20.2	60.8	373	10	DY074046	001015BEM
112	20.6	62.0	648	1	AA126476	zn85c10.s	C 185	20.2	60.8	381	4	C22446	Rice
113	20.6	62.0	649	14	AG944001	Drosophil	C 186	20.2	60.8	383	3	BQ494962	EST04191
114	20.6	62.0	652	14	DU811766	OG_Aba008	C 187	20.2	60.8	398	8	CV723952	root--05-
115	20.6	62.0	700	13	CZ056561	OM_Ba005	C 188	20.2	60.8	426	11	AQ056149	CIT-HSP-2
116	20.6	62.0	704	14	DU814390	OG_Aba010	C 189	20.2	60.8	444	1	AA485662	ab10dl1.s
117	20.6	62.0	712	13	CZ058391	OM_Ba005	C 190	20.2	60.8	450	5	C1176130	C1176130
118	20.6	62.0	752	14	AG390082	Mus_muscu	C 191	20.2	60.8	460	14	DX142283	OG_Aba017
119	20.6	62.0	765	14	DU648936	OG_Aba007	C 192	20.2	60.8	463	14	DU470222	109841550
120	20.6	62.0	769	14	DX151464	OG_Aba018	C 193	20.2	60.8	479	1	AJ669867	AJ669867
121	20.6	62.0	919	7	BE367854	601217593	C 194	20.2	60.8	495	4	CB363397	ZF001-P00
122	20.6	62.0	935	13	CL468062	SAIL_1280	C 195	20.2	60.8	500	11	AQ894773	HS 5489-A
123	20.6	62.0	998	1	AL536084	ALU536084	C 196	20.2	60.8	501	7	BF402331	UI-R-CA0-
124	20.6	62.0	1178	4	CA278162	SCBFSF203	C 197	20.2	60.8	510	11	AQ489635	RPCI-11-2
125	20.6	62.0	1181	12	CK214191	CH261-76K	C 198	20.2	60.8	514	5	C1082903	C1082903
126	20.4	61.4	334	14	BE192936	Branchios	C 199	20.2	60.8	525	12	BZ866263	CH240 237
127	20.4	61.4	375	5	CF065866	Ac3126 Am	C 200	20.2	60.8	543	14	DU820219	OG_Aba009
128	20.4	61.4	381	5	CF065857	Ac3118 Am	C 201	20.2	60.8	561	1	AJ819468	AJ819468
129	20.4	61.4	425	5	C1281861	C1281861	C 202	20.2	60.8	562	14	DX135946	DX135946
130	20.4	61.4	465	14	DX230415	OR_Aba011	C 203	20.2	60.8	584	14	DX137681	DX137681
131	20.4	61.4	469	14	DX199876	OR_Aba006	C 204	20.2	60.8	594	7	BF007745	BF007745
132	20.4	61.4	485	14	DX235652	OR_Aba011	C 205	20.2	60.8	595	14	DX148980	OG_Aba018
133	20.4	61.4	498	8	CV911310	PB025D1 m	C 206	20.2	60.8	606	14	DU813703	OG_Aba008
134	20.4	61.4	506	1	AL796157	AL796157	C 207	20.2	60.8	607	7	BB856878	BB856878
135	20.4	61.4	525	8	CV918677	PH011C1 c	C 208	20.2	60.8	612	7	BF529402	BF529402
136	20.4	61.4	544	12	CE491169	CE491169 tigr-gss-	C 209	20.2	60.8	620	13	CZ721912	OC_Ba004
137	20.4	61.4	551	4	CB996829	AGENCOURT	C 210	20.2	60.8	626	14	DX147831	OG_Aba018
138	20.4	61.4	577	14	DX202835	OR_Aba007	C 211	20.2	60.8	629	14	DX002808	OG_Aba012
139	20.4	61.4	585	5	CF985643	19767rsic	C 212	20.2	60.8	632	5	CF614469	CS008620
140	20.4	61.4	585	13	CL952531	OSIRUA001	C 213	20.2	60.8	648	9	DN247231	ACAE-aaa3
141	20.4	61.4	588	14	DX180616	OR_Aba004	C 214	20.2	60.8	648	10	DT886787	DT886787
142	20.4	61.4	596	5	CF333357	JMT--02-D	C 215	20.2	60.8	654	14	DX004311	OG_Aba010
143	20.4	61.4	603	14	DX330821	OR_Aba025	C 216	20.2	60.8	655	14	DX120462	OG_Aba014
144	20.4	61.4	625	14	AG017704	Homo sapi	C 217	20.2	60.8	657	13	DU293816	DU293816
145	20.4	61.4	641	14	DX252548	OR_Aba014	C 218	20.2	60.8	674	2	BI692846	60343149
146	20.4	61.4	666	14	DX279527	OR_Aba017	C 219	20.2	60.8	676	14	DU994749	OG_Aba011
147	20.4	61.4	684	14	DX350190	OR_Aba028	C 220	20.2	60.8	679	14	AG159835	Pan trogl
148	20.4	61.4	685	14	DX193912	OR_Aba006	C 221	20.2	60.8	681	13	DU152760	OG_Aba002
149	20.4	61.4	705	12	CE759798	tigr-gss-	C 222	20.2	60.8	708	5	CK321412	ACA_CNS_N
150	20.4	61.4	736	2	BJ601682	BJ601682	C 223	20.2	60.8	714	14	DX131596	OG_Aba016
151	20.4	61.4	738	14	DX340260	OR_Aba026	C 224	20.2	60.8	722	14	DX004586	OG_Aba010
152	20.4	61.4	742	13	DUI41749	OG_Aba004	C 225	20.2	60.8	730	13	DUI140578	OG_Aba001
153	20.4	61.4	745	14	DX342074	OG_Aba026	C 226	20.2	60.8	731	4	CA361232	CA361232
154	20.4	61.4	762	4	BX863913	BX863913	C 227	20.2	60.8	738	12	CE323040	CE323040
155	20.4	61.4	792	10	DT480584	WS02529.B	C 228	20.2	60.8	741	14	DU994526	tigr-gss-
156	20.4	61.4	841	13	CL923043	OA_Aba002	C 229	20.2	60.8	753	14	DU991909	OG_Aba010
157	20.4	61.4	849	10	DT485847	WS02529.B	C 230	20.2	60.8	756	14	DX123562	OG_Aba015
158	20.4	61.4	849	10	DV917461	CR	C 231	20.2	60.8	760	3	BW448787	BW448787
159	20.4	61.4	861	11	BZ506737	BONUA94TF	C 232	20.2	60.8	768	13	DUI64342	OG_Aba000
160	20.4	61.4	885	2	BG194062	RST13204	C 233	20.2	60.8	771	13	DUI57248	DUI57248
161	20.4	61.4	930	10	DV668338	GBI167 SS	C 234	20.2	60.8	782	13	DUI44355	OG_Aba003
162	20.4	61.4	1132	10	DV668323	GBI146 SS	C 235	20.2	60.8	785	14	DX144168	OG_Aba017
163	20.4	61.4	1253	9	DN712632	CNB03-E02	C 236	20.2	60.8	786	14	DX120765	OG_Aba014
164	20.4	61.4	1259	12	BZ579374	msh2 6326	C 237	20.2	60.8	789	13	DUI52353	OG_Aba003
165	20.2	60.8	226	1	AA653842	ns99905.s	C 238	20.2	60.8	793	14	DX141514	DX141514

239	20	60.2	803	14	DX166981	DX166981 OR_Aba002	C 312	19.8	59.6	881	7	BE570833	BE570833 601329334
240	20	60.2	807	14	DX131910	DX131910 OG_Aba016	313	19.8	59.6	881	14	AG896934	AG896934 Oryza sat
241	20	60.2	808	14	DX252243	DX252243 OR_Aba014	314	19.8	59.6	889	12	CG930068	CG930068 MBEIU05TF
c 242	20	60.2	808	14	CT347202	CT347202 Sus_scrof	315	19.8	59.6	902	8	CX118557	CX118557 EIL15K16
243	20	60.2	828	8	CO880092	CO880092 ZmGen.08	316	19.8	59.6	902	5	CK409360	CK409360 AUF_ifdvyr
c 244	20	60.2	841	8	CV626724	CV626724 zcm96C03	c 317	19.8	59.6	910	14	DX318579	DX318579 OR_Aba023
245	20	60.2	843	13	DUI163125	DUI163125 OG_Aba001	318	19.8	59.6	910	14	DX318579	DX318579 OR_Aba023
246	20	60.2	851	14	DX112393	DX112393 OG_Aba011	318	19.8	59.6	918	4	CB282262	CB282262 BT0233 B1
247	20	60.2	861	13	DUI140829	DUI140829 OG_Aba001	c 319	19.8	59.6	927	14	AG878261	AG878261 Oryza sat
248	20	60.2	870	14	DUI584647	DUI584647 OO_Ba007	c 320	19.8	59.6	1125	10	DT732936	DT732936 FGAS07885
c 249	20	60.2	875	13	DUI159652	DUI159652 OG_Aba002	c 321	19.8	59.6	1203	10	DT960076	DT960076 CFW166-G0
250	20	60.2	884	14	CT287810	CT287810 Sus_scrof	c 322	19.8	59.6	1236	9	DN674799	DN674799 CFW82-F06
251	20	60.2	888	13	DUI166307	DUI166307 OG_Aba002	323	19.8	59.6	1301	10	DT978691	DT978691 CLJ178-A1
252	20	60.2	889	14	DX120215	DX120215 OG_Aba014	324	19.8	59.6	1339	9	DN674268	DN674268 CFW79-D02
c 253	20	60.2	916	12	CG036134	CG036134 PUQ165TD	c 325	19.6	59.0	245	11	BH373947	BH373947 AG-ND-151
c 254	20	60.2	921	14	AG886656	AG886656 Oryza sat	c 326	19.6	59.0	258	13	CZ302987	CZ302987 ZMMBF0079
c 255	20	60.2	936	12	CG690881	CG690881 Oryza sat	c 327	19.6	59.0	262	8	CR453107	CR453107 CR453107
c 256	20	60.2	1005	12	CC279839	CC279839 Oryza sat	c 328	19.6	59.0	289	11	BH380315	BH380315 AG-ND-140
c 257	20	60.2	1108	9	DN734742	DN734742 CNB77-D10	c 329	19.6	59.0	317	1	AA413202	AA413202 ve98e01.r
c 258	20	60.2	1948	6	CNS0FY12	CR693466 Tetraodon	330	19.6	59.0	328	10	DV851330	DV851330 LH0224.CR
c 259	19.8	59.6	259	5	CK067009	CK067009 73577rsic	331	19.6	59.0	334	2	BF840932	BF840932 RC4-HT111
c 260	19.8	59.6	372	1	A1074739	A1074739 or83h08.s	c 332	19.6	59.0	360	13	CZ302982	CZ302982 ZMMBF0079
c 261	19.8	59.6	388	7	BF639640	BF639640 NF015C061	c 333	19.6	59.0	370	11	BH879815	BH879815 hc49a05.9
c 262	19.8	59.6	417	4	BY453405	BY453405 BY453405	334	19.6	59.0	371	13	CL892742	CL892742 abg15E12
c 263	19.8	59.6	420	1	AJ687003	AJ687003 AJ687003	c 335	19.6	59.0	480	13	CZ302984	CZ302984 ZMMBF0079
c 264	19.8	59.6	426	9	CX524301	CX524301 s13dnF14C	c 336	19.6	59.0	480	3	BP894128	BP894128 BP894128
c 265	19.8	59.6	438	11	AZ181494	AZ181494 SP_0184.B	c 337	19.6	59.0	490	11	BH374301	BH374301 AG-ND-169
c 266	19.8	59.6	459	1	AL370773	AL370773 MCB339H04	c 338	19.6	59.0	494	11	BH345437	BH345437 CH230-197
c 267	19.8	59.6	459	10	DT874937	DT874937 AGENCOURT	c 339	19.6	59.0	500	11	BH310403	BH310403 CH230-26F
c 268	19.8	59.6	477	1	AL368109	AL368109 MCB22C07	340	19.6	59.0	504	4	CA149658	CA149658 SCRCR2200
c 269	19.8	59.6	483	7	BE164659	BE164659 MUG011.A1	341	19.6	59.0	525	7	AW647923	AW647923 EST326377
c 270	19.8	59.6	484	7	BE205051	BE205051 EUGT397727	c 342	19.6	59.0	536	11	BH391241	BH391241 AG-ND-125
c 271	19.8	59.6	518	5	CK055939	CK055939 58641rsic	c 343	19.6	59.0	543	4	CA255066	CA255066 SCRPFL417
c 272	19.8	59.6	546	7	BF633119	BF633119 NF051C01D	344	19.6	59.0	545	4	BX359087	BX359087 BX359087
c 273	19.8	59.6	553	5	CK045300	CK045300 49524rsic	c 345	19.6	59.0	559	13	CZ302983	CZ302983 ZMMBF0079
c 274	19.8	59.6	558	14	AG017693	AG017693 Homo sapi	c 346	19.6	59.0	560	7	BE417682	BE417682 MUG023.D0
c 275	19.8	59.6	559	5	CK082488	CK082488 80396rsic	347	19.6	59.0	560	13	CHW600159	CHW600159 OA_Aba013
c 276	19.8	59.6	559	13	CW587038	CW587038 OA_Aba011	c 348	19.6	59.0	563	11	BH380402	BH380402 AG-ND-123
c 277	19.8	59.6	562	13	DUI196810	DUI196810 109855347	349	19.6	59.0	572	11	AQ982363	AQ982363 RPCI-23-3
c 278	19.8	59.6	566	4	CB638867	CB638867 OSJNEA07M	c 350	19.6	59.0	579	2	BM410582	BM410582 EST584909
c 279	19.8	59.6	566	8	CB986239	CB986239 PAQ3982.m	c 351	19.6	59.0	581	11	BH402515	BH402515 AG-ND-103
c 280	19.8	59.6	568	11	BZ155434	BZ155434 CH230-396	352	19.6	59.0	587	5	CD204272	CD204272 HS1_4.C06
c 281	19.8	59.6	574	13	CL412746	CL412746 RPCI144.43	353	19.6	59.0	591	12	CE040031	CE040031 T1gr-gss-
c 282	19.8	59.6	577	9	CX523993	CX523993 s13dnF12D	354	19.6	59.0	592	11	BZ210118	BZ210118 CH230-527
c 283	19.8	59.6	577	9	CX531318	CX531318 s13dnF04C	c 355	19.6	59.0	602	12	CG038368	CG038368 PUFWC83TB
c 284	19.8	59.6	598	4	CB622207	CB622207 OSI18A08K	c 356	19.6	59.0	602	13	CZ633978	CZ633978 OM_Ba017
c 285	19.8	59.6	598	11	AQ179932	AQ179932 HS_3203.B	c 357	19.6	59.0	606	9	DN875318	DN875318 nad38g08
c 286	19.8	59.6	601	10	DV753213	DV753213 PchrsEQ11	c 358	19.6	59.0	610	11	BH395274	BH395274 AG-ND-155
c 287	19.8	59.6	615	5	CK054047	CK054047 52848rsic	c 359	19.6	59.0	611	11	BH022354	BH022354 GH_MBB000
c 288	19.8	59.6	625	14	AG017694	AG017694 Homo sapi	360	19.6	59.0	613	1	AJ759050	AJ759050 AJ759050
c 289	19.8	59.6	631	7	BF646304	BF646304 NF074A12E	361	19.6	59.0	625	8	CO525273	CO525273 3530.1.16
c 290	19.8	59.6	636	2	B1267258	B1267258 NF104D091	c 362	19.6	59.0	625	5	CD207017	CD207017 HS1_26.H0
c 291	19.8	59.6	639	2	BJ823384	BJ823384 BJB23384	c 363	19.6	59.0	636	5	CD407079	CD407079 Gm_Ck3207
c 292	19.8	59.6	650	1	AM058624	AM058624 AM058624	364	19.6	59.0	640	2	BJ292648	BJ292648 BJ292648
c 293	19.8	59.6	652	2	BM488899	BM488899 pgm2n.pk0	c 365	19.6	59.0	640	8	CN660950	CN660950 A0712C04-
c 294	19.8	59.6	664	3	BQ148096	BQ148096 NF058F09F	c 366	19.6	59.0	647	13	CG370133	CG370133 f6b001f0
c 295	19.8	59.6	666	7	BF640986	BF640986 NF028E121	c 367	19.6	59.0	649	12	CG865094	CG865094 174B1T7.N
c 296	19.8	59.6	666	8	CV939041	CV939041 PNRpct.02	c 368	19.6	59.0	657	12	CG865092	CG865092 118F24T7
c 297	19.8	59.6	671	11	BH969963	BH969963 odj11f08.	c 369	19.6	59.0	659	5	CD233649	CD233649 SS1_3.A05
c 298	19.8	59.6	677	5	CK063722	CK063722 75069rsic	370	19.6	59.0	663	12	CL186923	CL186923 104_402.1
c 299	19.8	59.6	678	7	BF641301	BF641301 NF051B011	c 371	19.6	59.0	678	10	DR817803	DR817803 ZM_BFB005
c 300	19.8	59.6	700	13	CW586119	CW586119 OA_Aba011	c 372	19.6	59.0	685	5	CG935457	CG935457 OV_101J008
c 301	19.8	59.6	710	4	CB678615	CB678615 OSUNBE160	c 373	19.6	59.0	685	12	CG865076	CG865076 108E14T7
c 302	19.8	59.6	711	4	CB678615	CB678615 OSUNBE160	c 374	19.6	59.0	687	11	BH374178	BH374178 AG-ND-168
c 303	19.8	59.6	756	14	CT412996	CT412996 Sus_scrof	c 375	19.6	59.0	696	11	BH979878	BH979878 odg91h04.
c 304	19.8	59.6	769	2	BG581920	BG581920 EST483656	c 376	19.6	59.0	699	10	DV844402	DV844402 LH01219.C
c 305	19.8	59.6	781	8	CO163793	CO163793 FLD1_43.G	c 377	19.6	59.0	702	12	DV844402	DV844402 LH01219.C
c 306	19.8	59.6	815	14	DU583362	DU583362 OO_Ba007	c 378	19.6	59.0	710	3	BU938007	BU938007 AGENCOURT
c 307	19.8	59.6	818	14	DUA80901	DUA80901 109841593	c 379	19.6	59.0	712	11	BH384663	BH384663 AG-ND-138
c 308	19.8	59.6	839	4	CB631226	CB631226 OSIEB08K	c 380	19.6	59.0	714	5	CD935292	CD935292 OV_101D16
c 309	19.8	59.6	841	12	CG919708	CG919708 MBEUB73TF	c 381	19.6	59.0	714	13	CZ785482	CZ785482 OC_Ba014
c 310	19.8	59.6	862	14	AG861820	AG861820 Oryza sat	c 382	19.6	59.0	716	11	BZ260051	BZ260051 CH230-419
c 311	19.8	59.6	864	11	BZ249153	BZ249153 CH230-498	c 383	19.6	59.0	720	8	CO528946	CO528946 3530.1.19
							c 384	19.6	59.0	725	11	BH370133	BH370133 AG-ND-161

C 385	19.6	59.0	728	13	CW354811	FWB0001F0	458	19.4	58.4	461	7	BE502143	BE502143
C 386	19.6	59.0	730	3	BU349060	604171370	C 459	19.4	58.4	465	4	BE675014	BE675014
C 387	19.6	59.0	733	2	BU299093	BU299093	460	19.4	58.4	465	7	AW044486	AW044486
C 388	19.6	59.0	735	10	DV953868	SB03027B1	461	19.4	58.4	465	7	AW514161	AW514161
C 389	19.6	59.0	739	3	BU293942	603323255	462	19.4	58.4	467	14	DX235462	DX235462
C 390	19.6	59.0	740	1	AJ614305	AJ614305	C 463	19.4	58.4	469	3	BM852382	BM852382
C 391	19.6	59.0	743	7	BE415805	MUG001.CO	464	19.4	58.4	472	7	BE501526	BE501526
C 392	19.6	59.0	747	13	CL9111455	OA_Aba001	C 465	19.4	58.4	473	1	AJ038249	AJ038249
C 393	19.6	59.0	749	11	BZ003523	oeH91F09.	C 466	19.4	58.4	474	1	AA427486	AA427486
C 394	19.6	59.0	750	12	CG865087	116D149P6	467	19.4	58.4	475	1	AJ709346	AJ709346
C 395	19.6	59.0	751	14	AG393824	Mus muscu	C 468	19.4	58.4	476	14	DX196616	DX196616
C 396	19.6	59.0	752	11	BH387047	BH387047	469	19.4	58.4	476	14	DX306704	DX306704
C 397	19.6	59.0	757	8	CO518016	3530_1_11	C 470	19.4	58.4	477	1	AI118590	AI118590
C 398	19.6	59.0	759	12	CG283327	CG5DM367C	C 471	19.4	58.4	478	4	CA374622	CA374622
C 399	19.6	59.0	766	10	DR792922	ZM_BFB001	472	19.4	58.4	483	1	AI640610	AI640610
C 400	19.6	59.0	768	13	CZ808411	OC_Ba018	C 473	19.4	58.4	485	5	CK875130	CK875130
C 401	19.6	59.0	772	13	CZ271612	ZM_BFR0051	474	19.4	58.4	485	13	DUI148599	DUI148599
C 402	19.6	59.0	774	10	DR799211	ZM_BFB002	475	19.4	58.4	487	14	DX202775	DX202775
C 403	19.6	59.0	778	4	CA196818	SCBFA009	476	19.4	58.4	493	5	CF793715	CF793715
C 404	19.6	59.0	779	7	BE908400	601503023	477	19.4	58.4	496	1	AI082814	AI082814
C 405	19.6	59.0	783	10	DR965259	ZM_BFB008	478	19.4	58.4	497	1	AI672545	AI672545
C 406	19.6	59.0	784	2	BI829579	603079315	C 479	19.4	58.4	508	5	CD672812	CD672812
C 407	19.6	59.0	796	13	CW354812	FWB0001F0	C 480	19.4	58.4	513	1	AI623831	AI623831
C 408	19.6	59.0	809	3	BU227668	603797338	481	19.4	58.4	518	1	AA813357	AA813357
C 409	19.6	59.0	812	4	CB235964	AGENCOURT	482	19.4	58.4	518	1	AI190291	AI190291
C 410	19.6	59.0	815	7	BE194336	HVSMER008	C 483	19.4	58.4	518	5	CK866657	CK866657
C 411	19.6	59.0	816	3	BU119565	603142825	C 484	19.4	58.4	518	9	DB237826	DB237826
C 412	19.6	59.0	825	10	DV173400	ZM_BFB017	C 485	19.4	58.4	520	7	BE393690	BE393690
C 413	19.6	59.0	837	3	BU239087	603233967	486	19.4	58.4	520	14	DX233198	DX233198
C 414	19.6	59.0	840	11	BH403653	AG-ND-146	487	19.4	58.4	521	14	DX305420	DX305420
C 415	19.6	59.0	843	8	CN148798	WOUND1.58	488	19.4	58.4	528	10	DV951545	DV951545
C 416	19.6	59.0	852	10	DR806206	ZM_BFB003	489	19.4	58.4	528	11	BZ249649	BZ249649
C 417	19.6	59.0	872	13	CZ861141	OC_Ba025	C 490	19.4	58.4	528	14	DUI16826	DUI16826
C 418	19.6	59.0	874	13	CZ697321	OC_Ba001	C 491	19.4	58.4	530	14	DX227932	DX227932
C 419	19.6	59.0	882	8	CO815184	AGENCOURT	492	19.4	58.4	531	1	AI913001	AI913001
C 420	19.6	59.0	912	4	EX454774	EX454774	493	19.4	58.4	535	14	DX235689	DX235689
C 421	19.6	59.0	918	12	CG939142	MBELUS1RTR	C 494	19.4	58.4	538	1	AI652099	AI652099
C 422	19.6	59.0	935	14	DU736675	APK31883.	C 495	19.4	58.4	541	10	DT672790	DT672790
C 423	19.6	59.0	942	12	CC416069	FUHQG657D	C 496	19.4	58.4	541	11	AZ116714	AZ116714
C 424	19.6	59.0	948	10	DW612643	CLJ270-CO	C 497	19.4	58.4	550	13	CL920572	CL920572
C 425	19.6	59.0	966	14	CNS015SL	AL105855	C 498	19.4	58.4	550	14	DX263465	DX263465
C 426	19.6	59.0	992	13	DU092813	DUP02813	C 499	19.4	58.4	551	1	AJ707259	AJ707259
C 427	19.6	59.0	1023	14	CNS049WR	AL281124	C 500	19.4	58.4	552	2	BJ668716	BJ668716
C 428	19.6	59.0	1095	14	CNS012KO	AL101658	C 501	19.4	58.4	552	10	DT671238	DT671238
C 429	19.6	59.0	1155	9	DN733421	DN733421	C 502	19.4	58.4	555	12	CG812627	CG812627
C 430	19.6	59.0	1321	8	CV860693	CN870-A01	C 503	19.4	58.4	557	14	DX183948	DX183948
C 431	19.6	59.0	1464	2	BI458506	gonad_EST	C 504	19.4	58.4	558	14	DX202917	DX202917
C 432	19.6	59.0	2951	6	AK039029	BI458506	C 505	19.4	58.4	560	9	DA419705	DA419705
C 433	19.6	59.0	2952	6	AK138499	AK138499	C 506	19.4	58.4	563	2	BJ677748	BJ677748
C 434	19.6	59.0	4091	12	CG865108	phage3741	C 507	19.4	58.4	563	14	DX209586	DX209586
C 435	19.4	58.4	156	7	AW865253	PM4-SN002	C 508	19.4	58.4	580	4	CA382276	CA382276
C 436	19.4	58.4	218	8	CO28512	CKX028512	C 509	19.4	58.4	582	14	DX219250	DX219250
C 437	19.4	58.4	263	1	AI023526	AI023526	C 510	19.4	58.4	584	14	DX191898	DX191898
C 438	19.4	58.4	286	1	AA815461	AA815461	C 511	19.4	58.4	585	14	DX217929	DX217929
C 439	19.4	58.4	286	1	AA459692	AA459692	C 512	19.4	58.4	585	14	DX250186	DX250186
C 440	19.4	58.4	306	1	AA910794	AA910794	C 513	19.4	58.4	587	14	DUI12605	DUI12605
C 441	19.4	58.4	342	4	CB158087	K-EST0217	C 514	19.4	58.4	589	1	AI907844	AI907844
C 442	19.4	58.4	343	4	CA367145	642894_NC	C 515	19.4	58.4	590	1	AI872397	AI872397
C 443	19.4	58.4	359	9	CK624624	A24_A12.S	C 516	19.4	58.4	593	3	BQ009901	BQ009901
C 444	19.4	58.4	363	14	DU998011	OG_BBa013	C 517	19.4	58.4	594	3	BU632322	BU632322
C 445	19.4	58.4	364	10	DV535159	ZM_BFB022	C 518	19.4	58.4	595	14	DX223389	DX223389
C 446	19.4	58.4	367	1	AI005427	ou08c03.x	C 519	19.4	58.4	595	14	DX271711	DX271711
C 447	19.4	58.4	367	3	BP430822	BP430822	C 520	19.4	58.4	596	8	CN860235	CN860235
C 448	19.4	58.4	367	12	CE541719	tigr-gss-	C 521	19.4	58.4	601	8	CN261334	CN261334
C 449	19.4	58.4	368	11	AZ971882	ZM0245J02	C 522	19.4	58.4	601	14	DX246387	DX246387
C 450	19.4	58.4	376	3	BP431446	BP431446	C 523	19.4	58.4	604	10	DW249226	DW249226
C 451	19.4	58.4	402	13	DUI161690	DUI161690	C 524	19.4	58.4	606	13	BQ016866	BQ016866
C 452	19.4	58.4	411	14	DBU14274	OG_Aba010	C 525	19.4	58.4	608	14	DX174659	DX174659
C 453	19.4	58.4	416	1	AA609420	zu72C06.s	C 526	19.4	58.4	609	13	CW255829	CW255829
C 454	19.4	58.4	434	4	CA382669	662525_NC	C 527	19.4	58.4	612	14	DX244630	DX244630
C 455	19.4	58.4	435	4	CA329264	hab23e03.	C 528	19.4	58.4	612	14	DX244630	DX244630
C 456	19.4	58.4	439	8	CO515796	e13dSG54B	C 529	19.4	58.4	614	1	AI907902	AI907902
C 457	19.4	58.4	459	7	BF641113	NF032D031	C 530	19.4	58.4	616	14	DX273859	DX273859

531	19.4	58.4	618	8	CV666773	CV666773	LCPE04EX0	604	19.4	58.4	708	4	BX103042
532	19.4	58.4	619	4	BX862958	BX862958	BX862958	605	19.4	58.4	708	13	CV567782 OA_Aba009
533	19.4	58.4	619	4	CX427790	CX427790	UT-H-DF0-	606	19.4	58.4	711	2	BM476371
534	19.4	58.4	619	14	DX215811	DX215811	OR_Aba009	607	19.4	58.4	712	14	DU806698 OG_Aba009
535	19.4	58.4	622	14	DU817725	DU817725	OG_Aba008	608	19.4	58.4	712	3	AG180224 Pan trogl
536	19.4	58.4	623	2	BG576637	BG576637	602597521	609	19.4	58.4	713	3	BQ045183
537	19.4	58.4	623	13	DU144987	DU144987	OG_Aba001	610	19.4	58.4	714	8	CA446215 UI-H-ED1-
538	19.4	58.4	624	14	CT105224	CT105224	Sub scrof	611	19.4	58.4	718	13	CV651736 OA_Aba019
539	19.4	58.4	625	5	CD768360	CD768360	AGENCOURT	612	19.4	58.4	719	14	DU996741
540	19.4	58.4	625	14	DX210259	DX210259	OR_Aba008	613	19.4	58.4	724	14	DX294914 OR_Aba020
541	19.4	58.4	625	14	DX334457	DX334457	OR_Aba025	614	19.4	58.4	729	14	DX140198 OG_Aba017
542	19.4	58.4	626	4	BX305837	BX305837	BX305837	615	19.4	58.4	730	4	BX887640 BX887640
543	19.4	58.4	626	14	DX211261	DX211261	OR_Aba008	616	19.4	58.4	734	3	BU931024
544	19.4	58.4	626	14	DX313597	DX313597	OR_Aba022	617	19.4	58.4	737	2	BI084293 602869803
545	19.4	58.4	627	7	AW953769	AW953769	EST155839	618	19.4	58.4	737	8	CX259985 1313359 N
546	19.4	58.4	632	3	BM669571	BM669571	UI-E-DX1-	619	19.4	58.4	738	13	CZ160532 OR_Aba000
547	19.4	58.4	633	12	CC762984	CC762984	CH240_3G2	620	19.4	58.4	738	14	DX351450 OR_Aba028
548	19.4	58.4	635	14	DX277717	DX277717	OR_Aba017	621	19.4	58.4	743	13	DU140517 OG_Aba001
549	19.4	58.4	636	13	DUI62226	DUI62226	OG_Aba002	622	19.4	58.4	744	13	CK005755 ZNMBLa000
550	19.4	58.4	637	4	CA226603	CA226603	SCRLFL300	623	19.4	58.4	744	14	DX002693 OG_Aba012
551	19.4	58.4	637	9	CX526596	CX526596	el3dnp36R	624	19.4	58.4	747	14	DX136869
552	19.4	58.4	637	14	DU809403	DU809403	OG_Aba009	625	19.4	58.4	748	2	BI084624 602869803
553	19.4	58.4	638	11	AQ627943	AQ627943	CITBI-EI-	626	19.4	58.4	751	3	BU608408
554	19.4	58.4	639	14	DX277862	DX277862	OR_Aba017	627	19.4	58.4	753	10	DV945819 SB03003B1
555	19.4	58.4	640	14	DX196506	DX196506	OR_Aba006	628	19.4	58.4	753	14	DX252213 OR_Aba014
556	19.4	58.4	640	14	DX332282	DX332282	OR_Aba025	629	19.4	58.4	757	14	DX229940 OR_Aba010
557	19.4	58.4	641	11	BH371825	BH371825	AG-ND-167	630	19.4	58.4	757	14	DU649327 OG_Aba006
558	19.4	58.4	642	2	BI282053	BI282053	602953462	631	19.4	58.4	759	2	BG585200
559	19.4	58.4	642	13	CV545428	CV545428	OA_Aba006	632	19.4	58.4	762	5	CD367309 UI-H-FT2-
560	19.4	58.4	643	10	DV403860	DV403860	NADVP49TR	633	19.4	58.4	764	14	DX324466 OR_Aba024
561	19.4	58.4	643	14	DX008189	DX008189	OG_Aba011	634	19.4	58.4	768	14	DX316760 OR_Aba023
562	19.4	58.4	644	7	BE077683	BE077683	CM3-BT061	635	19.4	58.4	769	7	BF525703 60270033
563	19.4	58.4	645	2	BG284343	BG284343	602408446	636	19.4	58.4	769	14	DX315970 OR_Aba023
564	19.4	58.4	645	14	CV327872	CV327872	OR_Aba024	637	19.4	58.4	774	5	CK025650 AGENCOURT
565	19.4	58.4	646	3	BU625907	BU625907	UI-H-FG1-	638	19.4	58.4	775	1	AL576115 AU576115
566	19.4	58.4	647	14	DX274633	DX274633	OR_Aba017	639	19.4	58.4	775	14	BI262078 602953492
567	19.4	58.4	650	11	AQ521590	AQ521590	HS_5233_A	640	19.4	58.4	775	14	DX173222 OR_Aba003
568	19.4	58.4	650	14	DX265968	DX265968	OR_Aba016	641	19.4	58.4	779	12	CC091519 CSU-K33r-
569	19.4	58.4	651	14	DX005431	DX005431	OG_Aba010	642	19.4	58.4	780	10	AV721881 AV721881
570	19.4	58.4	651	14	DX346913	DX346913	OR_Aba027	643	19.4	58.4	780	10	AV721881
571	19.4	58.4	652	13	CV642605	CV642605	OA_Aba017	644	19.4	58.4	781	4	BX083439
572	19.4	58.4	655	14	DX170609	DX170609	OR_Aba002	645	19.4	58.4	785	14	DX115767
573	19.4	58.4	657	4	CA346536	CA346536	677436 NC	646	19.4	58.4	785	14	DX296780 OR_Aba020
574	19.4	58.4	658	13	CV589467	CV589467	OA_Aba012	647	19.4	58.4	787	13	DUI55819 OG_Aba003
575	19.4	58.4	658	14	DX128903	DX128903	OG_Aba015	648	19.4	58.4	791	14	DX249524 OR_Aba013
576	19.4	58.4	658	14	CV350527	CV350527	OR_Aba028	649	19.4	58.4	795	14	DX249524 OR_Aba007
577	19.4	58.4	659	13	CT759406	CT759406	OC_Ba011	650	19.4	58.4	798	11	BZ208571 CH230-342
578	19.4	58.4	659	13	CZ780880	CZ780880	OC_Ba014	651	19.4	58.4	798	14	DX121208 OG_Aba014
579	19.4	58.4	663	8	CO128674	CO128674	GR_Eb25K	652	19.4	58.4	798	14	DX298054 OR_Aba020
580	19.4	58.4	663	13	CZ497072	CZ497072	OA_BBa013	653	19.4	58.4	799	10	DT600470 km103-4ms
581	19.4	58.4	664	1	AI133130	AI133130	HA1744_Hu	654	19.4	58.4	801	14	DT600470 km103-4ms
582	19.4	58.4	665	8	CX247445	CX247445	1299342 N	655	19.4	58.4	804	10	DT603023 km101-23m
583	19.4	58.4	665	14	DX190129	DX190129	OR_Aba005	656	19.4	58.4	805	13	DUI39117 OG_Aba002
584	19.4	58.4	667	13	CZ184596	CZ184596	OA_BBa010	657	19.4	58.4	805	13	DUI52581 OG_Aba003
585	19.4	58.4	668	7	BE077702	BE077702	CM3-BT061	658	19.4	58.4	805	14	DUI663276 OG_Aba005
586	19.4	58.4	670	8	CX038914	CX038914	1352007 N	659	19.4	58.4	807	10	DW557947 EST_ssal
587	19.4	58.4	670	9	CX718756	CX718756	1335363 N	660	19.4	58.4	807	14	DX300126 OR_Aba020
588	19.4	58.4	673	14	DX111972	DX111972	OG_Aba013	661	19.4	58.4	810	14	DX178333 OR_Aba003
589	19.4	58.4	675	14	DX229298	DX229298	OR_Aba010	662	19.4	58.4	810	14	CT1179196 Sub scrof
590	19.4	58.4	678	13	CV311420	CV311420	104_800_1	663	19.4	58.4	812	14	DUI39117 OG_Aba002
591	19.4	58.4	678	14	DX220448	DX220448	OR_Aba009	664	19.4	58.4	812	14	DUI663418 OG_Aba007
592	19.4	58.4	679	14	DX324186	DX324186	ABa024	665	19.4	58.4	814	14	DX133780 OG_Aba016
593	19.4	58.4	681	13	CV617140	CV617140	OA_Aba016	666	19.4	58.4	821	13	CZ703729 OC_Ba002
594	19.4	58.4	682	14	DX189871	DX189871	OR_Aba005	667	19.4	58.4	823	13	DUI65177 OG_Aba002
595	19.4	58.4	683	13	CL925326	CL925326	OA_Aba005	668	19.4	58.4	831	14	DX127199 OG_Aba015
596	19.4	58.4	688	14	AG296055	AG296055	MuB_muscu	669	19.4	58.4	833	2	BG425683 602452991
597	19.4	58.4	688	14	AG296055	AG296055	MuB_muscu	670	19.4	58.4	835	2	BI836621 603089326
598	19.4	58.4	698	14	CA369452	CA369452	645908 NC	671	19.4	58.4	836	12	BZ843024 CH240_249
599	19.4	58.4	701	4	DX277007	DX277007	OR_Aba017	672	19.4	58.4	855	14	DX131535 OG_Aba016
600	19.4	58.4	703	8	CO398610	CO398610	AGENCOURT	673	19.4	58.4	867	13	DUI57918 OG_Aba003
601	19.4	58.4	704	14	DX663496	DX663496	OG_Aba007	674	19.4	58.4	869	14	BU957142 AGENCOURT
602	19.4	58.4	706	14	DX154664	DX154664	OR_Aba000	675	19.4	58.4	874	4	BX324875 BX324875
603	19.4	58.4	708	3	BM975880	BM975880	UI-CF-EN1	676	19.4	58.4	879	14	DU801832 OG_Aba007

C 677	19.4	58.4	880	11	AQ574193	AQ574193 nbxb00830	750	19.2	57.8	523	12	B2821846	B2821846
C 678	19.4	58.4	880	13	DUI60263	DUI60263 OG_Aba002	751	19.2	57.8	526	2	B112783	B112783
C 679	19.4	58.4	892	4	BX392781	BX392781 BX392781	C 752	19.2	57.8	527	3	BQ270361	BQ270361
C 680	19.4	58.4	901	4	BX351273	BX351273 BX351273	753	19.2	57.8	528	2	BG967089	BG967089
C 681	19.4	58.4	902	14	DX248750	DX248750 OR_Aba013	754	19.2	57.8	534	11	AQ519437	AQ519437
C 682	19.4	58.4	903	10	DR571159	DR571159 WS00728.B	C 755	19.2	57.8	535	2	BJ137179	BJ137179
C 683	19.4	58.4	905	14	CNS052NY	AL18391 Tetraodon	C 756	19.2	57.8	539	14	DX270940	DX270940
C 684	19.4	58.4	906	14	DX154493	DX154493 OR_Aba000	C 757	19.2	57.8	544	3	BQ572270	BQ572270
C 685	19.4	58.4	914	13	C2930239	C2930239 247871 TO	C 758	19.2	57.8	548	5	CF312148	CF312148
C 686	19.4	58.4	916	8	CV661456	CV661456 LCPE04EX0	759	19.2	57.8	555	13	C2765143	C2765143
C 687	19.4	58.4	918	8	CV661456	BQ670244 AGENCOURT	C 760	19.2	57.8	567	14	DX290429	DX290429
C 688	19.4	58.4	920	3	BQ670244	BQ670244 AGENCOURT	C 761	19.2	57.8	569	9	DB090514	DB090514
C 689	19.4	58.4	921	3	BQ672967	BQ672967 AGENCOURT	C 762	19.2	57.8	579	9	DB090542	DB090542
C 690	19.4	58.4	930	3	BQ672967	BQ672967 AGENCOURT	C 763	19.2	57.8	583	3	BU579107	BU579107
C 691	19.4	58.4	937	13	DUI56265	DUI56265 OG_Aba003	764	19.2	57.8	593	11	AQ9833627	AQ9833627
C 692	19.4	58.4	948	14	CNS028GB	AL189812 Tetraodon	C 765	19.2	57.8	598	11	BH033031	BH033031
C 693	19.4	58.4	958	7	BE873351	BE873351 601450410	C 766	19.2	57.8	610	7	BF126257	BF126257
C 694	19.4	58.4	995	12	CL100808	CL100808 ISB1_36G1	C 767	19.2	57.8	613	8	CO984040	CO984040
C 695	19.4	58.4	1002	1	AL578908	AL578908 AL578908	C 768	19.2	57.8	621	14	DE044953	DE044953
C 696	19.4	58.4	1008	3	BQ673587	BQ673587 AGENCOURT	C 769	19.2	57.8	626	1	AL710408	AL710408
C 697	19.4	58.4	1015	3	BU838657	BU838657 AGENCOURT	C 770	19.2	57.8	628	6	CNS0F18J	CR676871 Tetraodon
C 698	19.4	58.4	1037	1	AL572304	AL572304 AL572304	771	19.2	57.8	630	6	CNS0F80L	CR660591 Tetraodon
C 699	19.4	58.4	1072	4	BX343804	BX343804 BX343804	772	19.2	57.8	631	6	CNS0F8AO	CR662682 Tetraodon
C 700	19.4	58.4	1090	1	AL573788	AL573788 AL573788	773	19.2	57.8	643	12	CC431822	CC431822
C 701	19.4	58.4	1097	7	BF108105	BF108105 601824166	C 774	19.2	57.8	643	14	CNS03AJH	AL235286 Tetraodon
C 702	19.4	58.4	1108	2	BG745887	BG745887 602724058	C 775	19.2	57.8	647	13	CM420322	CM420322
C 703	19.4	58.4	1115	1	AL582294	AL582294 AL582294	C 776	19.2	57.8	652	13	C2761776	CF663081 CGLMO9a34
C 704	19.4	58.4	1289	14	AG320739	AG320739 Mus muscu	777	19.2	57.8	654	5	CF663081	CF663081
C 705	19.4	58.4	1962	13	CL946413	CL946413 OsIFSB004	778	19.2	57.8	656	11	B56756	B56756
C 706	19.4	58.4	2214	6	CR595335	CR595335 full-leg	C 779	19.2	57.8	656	11	B56756	B56756
C 707	19.4	58.4	2321	6	HSW801772	AL136804 Homo sapi	C 780	19.2	57.8	690	14	DX244669	DX244669
C 708	19.4	58.4	4471	6	AK134594	AK134594 Mus muscu	781	19.2	57.8	696	12	CE752046	CE752046
C 709	19.2	57.8	186	11	A2310055	AZ310055 IM0018A13	782	19.2	57.8	696	13	CM829026	CM829026
C 710	19.2	57.8	210	7	BE703039	BE703039 QV4-NN110	783	19.2	57.8	702	7	BE303650	BE303650
C 711	19.2	57.8	222	10	H33336	H33336 EST109223 R	C 784	19.2	57.8	703	5	CF477539	CF477539
C 712	19.2	57.8	237	4	CB095985	CB095985 ie98e10.b	C 785	19.2	57.8	705	13	C2785549	C2785549
C 713	19.2	57.8	244	7	BF566325	BF566325 UI-R-BT1	C 786	19.2	57.8	709	4	BY741333	BY741333
C 714	19.2	57.8	257	9	DNI99131	DNI99131 USDA-FP 1	787	19.2	57.8	711	5	CK471842	CK471842
C 715	19.2	57.8	299	10	CV366014	CV366014 ZO_Ee000	C 788	19.2	57.8	711	14	AG323047	AG323047
C 716	19.2	57.8	312	8	CV300161	CV300161 EST889A08	789	19.2	57.8	716	12	CE565040	CE565040
C 717	19.2	57.8	326	1	AA281700	AA281700 zt03e01.r	C 790	19.2	57.8	721	12	B2713884	B2713884
C 718	19.2	57.8	336	1	AA282102	AA282102 zt03e01.r	C 791	19.2	57.8	721	12	B2713872	B2713872
C 719	19.2	57.8	338	1	AA083796	AA083796 zmg3c01.r	C 792	19.2	57.8	730	14	DU824383	DU824383
C 720	19.2	57.8	343	5	CO643720	CO643720 USDA-FP 1	C 793	19.2	57.8	734	4	CB951670	CB951670
C 721	19.2	57.8	364	5	CF119241	CF119241 MTU10CS.P	C 794	19.2	57.8	734	12	CG358978	CG358978
C 722	19.2	57.8	367	1	AA112374	AA112374 zn68e10.r	C 795	19.2	57.8	734	14	CT267185	CT267185
C 723	19.2	57.8	369	10	N59359	N59359 yv69f07.bl	C 796	19.2	57.8	737	7	BF503588	BF503588
C 724	19.2	57.8	372	10	T09412	T09412 EST07305 In	C 797	19.2	57.8	739	12	CE355967	CE355967
C 725	19.2	57.8	400	10	H66744	H66744 yr83h09.bl	C 798	19.2	57.8	740	11	AQ269767	AQ269767
C 726	19.2	57.8	401	4	CB698881	CB698881 AMGNNUC:S	C 799	19.2	57.8	740	13	CM420423	CM420423
C 727	19.2	57.8	404	4	CB811493	CB811493 AMGNNUC:N	C 800	19.2	57.8	746	12	CC407518	CC407518
C 728	19.2	57.8	405	4	CA106458	CA106458 SCQGH101	C 801	19.2	57.8	749	8	CO400087	CO400087
C 729	19.2	57.8	405	4	CA108448	CA108448 SCSPHR104	C 802	19.2	57.8	761	3	BU418329	BU418329
C 730	19.2	57.8	406	4	CB007662	CB007662 AMGNNUC:S	C 803	19.2	57.8	762	2	BM292365	BM292365
C 731	19.2	57.8	406	14	DE276296	DE276296 Oryzias 1	C 804	19.2	57.8	762	14	DU575682	DU575682
C 732	19.2	57.8	408	5	C1315825	C1315825 C1315825	C 805	19.2	57.8	765	7	BF505955	BF505955
C 733	19.2	57.8	414	8	CV359747	CV359747 PMO-FT005	C 806	19.2	57.8	770	12	CC733202	CC733202
C 734	19.2	57.8	415	4	CB800147	CB800147 AMGNNUC:S	C 807	19.2	57.8	773	3	BP161116	BP161116
C 735	19.2	57.8	419	1	AA136770	AA136770 zn96b05.r	C 808	19.2	57.8	778	14	CT358500	CT358500
C 736	19.2	57.8	419	11	AQ132041	AQ132041 HS_2170.A	C 809	19.2	57.8	790	9	DN936415	DN936415
C 737	19.2	57.8	428	11	HS8662	HS8662 HS_1031-A2	C 810	19.2	57.8	794	12	B2821848	B2821848
C 738	19.2	57.8	429	13	CL811987	CL811987 OR_Cha002	C 811	19.2	57.8	798	7	BE573112	BE573112
C 739	19.2	57.8	430	5	CK101956	CK101956 G058P16.5	C 812	19.2	57.8	798	12	CG449214	CG449214
C 740	19.2	57.8	430	7	AW974845	AW974845 EST386950	C 813	19.2	57.8	799	8	CK242254	CK242254
C 741	19.2	57.8	433	11	A2893300	A2893300 RPT1-24-2	C 814	19.2	57.8	805	12	CC360740	CC360740
C 742	19.2	57.8	433	11	AQ221389	AQ221389 HS_2004.B	C 815	19.2	57.8	806	14	DX328006	DX328006
C 743	19.2	57.8	435	4	CB792480	CB792480 AMGNNUC:N	C 816	19.2	57.8	808	11	BH355348	BH355348
C 744	19.2	57.8	435	12	CE668852	CE668852 tigr-gss-	C 817	19.2	57.8	812	12	CG015910	CG015910
C 745	19.2	57.8	455	4	CB741393	CB741393 AMGNNUC:N	C 818	19.2	57.8	812	12	CC059257	CC059257
C 746	19.2	57.8	504	13	CZ600597	CZ600597 OM_Ba009	C 819	19.2	57.8	818	11	AQ686559	AQ686559
C 747	19.2	57.8	512	5	CK575086	CK575086 IST_W15.8	C 820	19.2	57.8	822	12	B2824770	B2824770
C 748	19.2	57.8	513	3	BP488415	BP488415 BP488415	C 821	19.2	57.8	834	14	DU598438	DU598438
C 749	19.2	57.8	515	11	AZ829820	AZ829820 2M0107M08	C 822	19.2	57.8	838	14	DU589035	DU589035

823	19.2	57.8	839	14	CTI139326	CTI139326 Sus scrofa	C 896	19	57.2	448	4	CB745984	CB745984 AMGNNUC:N
824	19.2	57.8	841	11	BH116290	BH116290 RPI-24-3	C 897	19	57.2	450	2	BG054787	BG054787 nac9f02.
825	19.2	57.8	843	8	CV073992	CV073992 AGENCOURT	C 898	19	57.2	452	3	BQ160084	BQ160084 WHE1262.H
826	19.2	57.8	843	12	CC520254	CC520254 CH240.367	C 899	19	57.2	453	3	BQ657485	BQ657485 HA08C19U
827	19.2	57.8	850	2	BI666483	BI666483 603288769	C 900	19	57.2	453	4	CB785195	CB785195 AMGNNUC:N
828	19.2	57.8	856	12	CL136218	CL136218 ISB1-107K	C 901	19	57.2	453	4	DB218013	DB218013 DB218013
829	19.2	57.8	858	13	CV784431	CV784431 SP_Bb001	C 902	19	57.2	460	14	DX005731	DX005731 OG_ABA012
830	19.2	57.8	859	4	CA764261	CA764261 AF53-Rpf	C 903	19	57.2	468	2	BG673649	BG673649 DRNDE10
831	19.2	57.8	859	14	CTI190881	CTI190881 Sus scrofa	C 904	19	57.2	469	2	BG672041	BG672041 DRNCBC01
832	19.2	57.8	860	14	AG838771	AG838771 Oryza sat	C 905	19	57.2	472	11	AQ860313	AQ860313 nbeab0014N
833	19.2	57.8	864	11	AQ576152	AQ576152 nxb00088N	C 906	19	57.2	472	14	DX109080	DX109080 OG_ABA013
834	19.2	57.8	869	12	BZ824774	BZ824774 PUGC260TD	C 907	19	57.2	473	11	AQ924279	AQ924279 RPI-23-2
835	19.2	57.8	871	14	AG906273	AG906273 Oryza sat	C 908	19	57.2	475	14	DX293321	DX293321 OR_ABA019
836	19.2	57.8	873	12	CG287047	CG287047 QG5BN52TC	C 909	19	57.2	476	1	AL699869	AL699869 DKFp666P
837	19.2	57.8	874	14	CNS03JYF	AL247488 Tetraodon	C 910	19	57.2	482	2	BI471948	BI471948 sal97e08
838	19.2	57.8	886	14	CT211216	CT211216 Sus scrofa	C 911	19	57.2	488	9	DB225351	DB225351 DB225351
839	19.2	57.8	891	8	CO402600	CO402600 AGENCOURT	C 912	19	57.2	491	11	AZ127915	AZ127915 OJUNB0009
840	19.2	57.8	903	7	BF539190	BF539190 602053054	C 913	19	57.2	496	5	CF076262	CF076262 QHKN10N20.
841	19.2	57.8	908	12	CC598822	CC598822 ZMWBBC039	C 914	19	57.2	497	2	BG345699	BG345699 dg40a09.Y
842	19.2	57.8	916	3	BU523575	BU523575 AGENCOURT	C 915	19	57.2	499	4	BW890887	BW890887 BW890887
843	19.2	57.8	920	12	CC355380	CC355380 PUHE193TB	C 916	19	57.2	499	13	CZ432972	CZ432972 OA_BBA012
844	19.2	57.8	920	12	CG230459	CG230459 CGWHE17TH	C 917	19	57.2	499	13	DUI65498	DUI65498 OG_ABA002
845	19.2	57.8	921	8	CV237396	CV237396 WS01227.B	C 918	19	57.2	500	7	BF220768	BF220768 NXCI_150
846	19.2	57.8	932	14	CNS0446J	AL273700 Tetraodon	C 919	19	57.2	500	14	CT087968	CT087968 Sus scrofa
847	19.2	57.8	942	6	CNS0FTD7	CR687423 Tetraodon	C 920	19	57.2	507	14	DX002489	DX002489 OG_ABA012
848	19.2	57.8	942	6	CNS0G4ET	CR701737 Tetraodon	C 921	19	57.2	507	14	DX142095	DX142095 OG_ABA017
849	19.2	57.8	942	12	CL113276	CL113276 ISB1-58E2	C 922	19	57.2	507	14	DX234697	DX234697 OR_ABA011
850	19.2	57.8	943	12	CC365222	CC365222 PUHRE44TB	C 923	19	57.2	509	9	DA756322	DA756322 DA756322
851	19.2	57.8	944	4	CB570015	CB570015 AGENCOURT	C 924	19	57.2	509	14	DX233314	DX233314 OR_ABA011
852	19.2	57.8	945	12	CG110398	CG110398 PUIGC94TD	C 925	19	57.2	510	11	B24496	B24496 F20H15TF IG
853	19.2	57.8	946	2	BI76067	BG176067 602335293	C 926	19	57.2	512	9	DA413663	DA413663 DA413663
854	19.2	57.8	948	4	BX928536	BX928536 BX928536	C 927	19	57.2	514	9	DA756918	DA756918 DA756918
855	19.2	57.8	948	6	CNS0PAOS	CP663190 Tetraodon	C 928	19	57.2	518	9	DB236974	DB236974 DB236974
856	19.2	57.8	950	2	BI685093	BI685093 603310070	C 929	19	57.2	520	3	BW046535	BW046535 BW046535
857	19.2	57.8	957	12	CG059256	CG059256 PUFN164TB	C 930	19	57.2	526	3	BP033627	BP033627 BP033627
858	19.2	57.8	968	13	CL792500	CL792500 EST663861	C 931	19	57.2	528	13	DUI60723	DUI60723 OG_ABA000
859	19.2	57.8	981	13	CL945505	CL945505 OeIFSB004	C 932	19	57.2	531	9	DB132388	DB132388 DB132388
860	19.2	57.8	983	13	CL242026	CL242026 ZMWBBC058	C 933	19	57.2	532	9	DA215884	DA215884 DA120033
861	19.2	57.8	988	6	CNS0GLG4	CR723822 Tetraodon	C 934	19	57.2	533	9	DA120033	DA120033 DA120033
862	19.2	57.8	1037	13	CV923681	CV923681 EDCA294TR	C 935	19	57.2	533	9	DA807619	DA807619 DA807619
863	19.2	57.8	1045	14	DU748934	DU748934 ASNC806.B	C 936	19	57.2	534	9	DA106636	DA106636 DA106636
864	19.2	57.8	1085	2	BM541329	BM541329 AGENCOURT	C 937	19	57.2	534	9	DB001353	DB001353 DB001353
865	19.2	57.8	1100	7	BE367710	BE367710 601217495	C 938	19	57.2	534	9	DB271983	DB271983 DB271983
866	19.2	57.8	1245	13	CL980863	CL980863 OeIFCC045	C 939	19	57.2	535	2	BG461762	BG461762 RST44674
867	19.2	57.8	1567	6	AK148211	AK148211 Mub mubcu	C 940	19	57.2	535	9	DB065211	DB065211 DB065211
868	19	57.2	167	10	DA224260	DA224260 EST-AR154	C 941	19	57.2	536	7	BF282165	BF282165 EST446756
869	19	57.2	161	1	AI582072	AI582072 ar96a11.x	C 942	19	57.2	536	7	AW916026	AW916026 EST473730
870	19	57.2	170	3	BP726758	BP726758 BP726758	C 943	19	57.2	536	9	DB002518	DB002518 DB002518
871	19	57.2	186	6	DQ083178	DQ083178 Oryza sat	C 944	19	57.2	536	14	DUI993951	DUI993951 OG_ABA011
872	19	57.2	209	4	BX474430	BX474430 DKF2p686M	C 945	19	57.2	536	14	DX131129	DX131129 OG_ABA016
873	19	57.2	228	10	DW443982	DW443982 HHAGE0417	C 946	19	57.2	537	9	DA856503	DA856503 DA856503
874	19	57.2	258	13	CL436972	CL436972 FST4243-N	C 947	19	57.2	538	2	BU294687	BU294687 BU294687
875	19	57.2	265	1	AV412994	AV412994 AV412994	C 948	19	57.2	539	2	BF737000	BF737000 PM4-KT004
876	19	57.2	276	7	BB416586	BB416586 BB416586	C 949	19	57.2	539	9	DA096945	DA096945 DA096945
877	19	57.2	294	14	DX151594	DX151594 OG_ABA018	C 950	19	57.2	539	9	DA187652	DA187652 DA187652
878	19	57.2	310	2	BG094915	BG094915 mac25e12.	C 951	19	57.2	540	7	BB772692	BB772692 BB772692
879	19	57.2	321	2	BG941396	BG941396 ax12d12.x	C 952	19	57.2	540	9	DA054122	DA054122 DA054122
880	19	57.2	323	7	BB240065	BB240065 BB240065	C 953	19	57.2	542	9	DA187665	DA187665 DA187665
881	19	57.2	325	7	BB105464	BB105464 BB105464	C 954	19	57.2	543	8	CN846000	CN846000 PG07006F0
882	19	57.2	350	1	AI237259	AI237259 EST233821	C 955	19	57.2	544	9	DA379405	DA379405 DA379405
883	19	57.2	352	4	CA004729	CA004729 HSL103R	C 956	19	57.2	545	4	BW709331	BW709331 BW709331
884	19	57.2	381	7	AW928533	AW928533 EST337321	C 957	19	57.2	545	9	DB213625	DB213625 DB213625
885	19	57.2	391	4	BY303775	BY303775 BY303775	C 958	19	57.2	546	5	CD923282	CD923282 G750.107J
886	19	57.2	391	13	CZ128776	CZ128776 OA_BBA001	C 959	19	57.2	546	9	DA314693	DA314693 DA314693
887	19	57.2	396	11	B52737	B52737 CIT-HSP-200	C 960	19	57.2	546	9	DB075803	DB075803 DB075803
888	19	57.2	401	11	AQ775802	AQ775802 HS_2006.A	C 961	19	57.2	546	9	DB087774	DB087774 DB087774
889	19	57.2	408	10	N44923	N44923 YV34G09.r1	C 962	19	57.2	547	9	DA402764	DA402764 DA402764
890	19	57.2	413	9	DA756329	DA756329 DA756329	C 963	19	57.2	547	9	DA955346	DA955346 DA955346
891	19	57.2	426	13	DUI61294	DUI61294 OG_ABA000	C 964	19	57.2	547	9	DB077018	DB077018 DB077018
892	19	57.2	438	1	AJ434048	AJ434048 AJ434048	C 965	19	57.2	548	9	DB138053	DB138053 DB138053
893	19	57.2	439	13	CZ100000	CZ100000 OM_Ba010	C 966	19	57.2	548	9	DB082782	DB082782 DB082782
894	19	57.2	446	4	CB744155	CB744155 AMGNNUC:N	C 967	19	57.2	549	9	DB285387	DB285387 DB285387
895	19	57.2	447	1	AA553125	AA553125 vJ51d11.r	C 968	19	57.2	550	9	DA211424	DA211424 DA211424

```

c 969      19 57.2      550 9 DA742945
c 970      19 57.2      550 9 DB219745
c 971      19 57.2      550 9 DB285983
c 972      19 57.2      551 9 DA800048
c 973      19 57.2      551 12 CC559317
c 974      19 57.2      552 9 DA652161
c 975      19 57.2      553 9 DB077106
c 976      19 57.2      553 9 DB237760
c 977      19 57.2      554 9 DA860830
c 978      19 57.2      554 11 A2718342
c 979      19 57.2      556 9 DA222240
c 980      19 57.2      556 9 DA406387
c 981      19 57.2      557 9 DA242298
c 982      19 57.2      558 9 DA117988
c 983      19 57.2      558 9 DB165266
c 984      19 57.2      558 10 DT026877
c 985      19 57.2      559 12 CE447270
c 986      19 57.2      560 9 DA757655
c 987      19 57.2      561 9 DA716058
c 988      19 57.2      562 8 CV913151
c 989      19 57.2      562 9 DA726057
c 990      19 57.2      563 9 DA587461
c 991      19 57.2      563 9 DA721371
c 992      19 57.2      564 9 DA720857
c 993      19 57.2      564 9 DA724206
c 994      19 57.2      565 3 BU473020
c 995      19 57.2      565 9 DA725571
c 996      19 57.2      565 9 DB123941
c 997      19 57.2      565 13 CZ035288
c 998      19 57.2      565 14 DX138293
c 999      19 57.2      566 9 DA460877
c1000     19 57.2      566 9 DA525595

ALIGNMENTS

RESULT 1
BI063311      158 bp      mRNA      linear      EST 15-JUN-2001
LOCUS      IL3-UT0117-300301-538-E01 UT0117 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BI063311
ACCESSION      BI063311.1 GI:14470838
VERSION      EST.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1 (bases 1 to 158)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t=IL3-UT0117-
300301-538-E01&t3=2001-03-30&t4=1)
Seq primer: puc 18 forward

FEATURES
source
High quality sequence stop: 158.
Location/Qualifiers
1..158
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UT0117"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESPEC PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 90.4%; Score 30; DB 2; Length 158;
Best Local Similarity 64.3%; Pred. No. 0.049;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:
Db 100 GCCACGCTTGCCGAGCCCTTACAAGACTATGAAGTAAAG 141

RESULT 2
BI063311/c
LOCUS      BE818833 307 bp      mRNA      linear      EST 21-SEP-2000
DEFINITION      CM2-BN0302-050700-256-c01 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BE818833
VERSION      BE818833.1 GI:10251067
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1 (bases 1 to 307)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2-BN0302-050
700-256-c01&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 307.

FEATURES
source
High quality sequence stop: 307.
Location/Qualifiers
1..307
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products

```

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector: Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 307;
Best Local Similarity 64.3%; Pred. No. 0.056; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 242 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 201

RESULT 3

BE818817/c
LOCUS BE818817 341 bp mRNA linear EST 21-SEP-2000
DEFINITION CM2-BN0302-050700-256-b07 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE818817
VERSION BE818817.1 GI:10251051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 341)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-050700-256-b07&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 341.

Location/Qualifiers
1. 341
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source
Query Match 90.4%; Score 30; DB 7; Length 341;
Best Local Similarity 64.3%; Pred. No. 0.058; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 242 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 201

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

RESULT 5
BE818828/c
LOCUS BE818828 396 bp mRNA linear EST 21-SEP-2000
DEFINITION CM2-BN0302-050700-256-a05 BN0302 Homo sapiens cDNA, mRNA sequence.

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 253 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 212

RESULT 4

BE818835/c
LOCUS BE818835 384 bp mRNA linear EST 21-SEP-2000
DEFINITION CM2-BN0302-050700-256-c11 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE818835
VERSION BE818835.1 GI:10251069
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-050700-256-c11&t3=2000-07-05&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 384.

Location/Qualifiers
1. 384
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source
Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 384;
Best Local Similarity 64.3%; Pred. No. 0.059; 8; Indels 0; Gaps 0;
Matches 27; Conservative 7; Mismatches 0;

Qy 1 GCNCAYCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:
Db 257 GCCCAGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 216

```
ACCESSION BE818828
VERSION BE818828.1 GI:10251062
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 396)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM2-BN0302-050
700-256-a05&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 396.
FEATURES
Location/Qualifiers
1..396
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 90.4%; Score 30; DB 7; Length 396;
Best Local Similarity 64.3%; Pred. No. 0.059;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAVCGTGGCCNTCCNTAYARGAYTAGGTAAAR 42
Db 248 GCCACGCTTGCCGCGCCCTTACAAAGACTATGAAGTAAAG 207
RESULT 6
BE818840/c
LOCUS CM2-BN0302-050700-256-f02 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818840
ACCESSION BE818840.1 GI:10251074
VERSION BE818840.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 404)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
```

```
ACCESSION BE818805
VERSION BE818805.1 GI:10251039
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 424)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
```

```
ACCESSION BE818805/c
LOCUS CM2-BN0302-040700-253-b10 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818805
ACCESSION BE818805.1 GI:10251039
VERSION BE818805.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 424)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
```

```
ACCESSION BE818805/c
LOCUS CM2-BN0302-040700-253-b10 BN0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE818805
ACCESSION BE818805.1 GI:10251039
VERSION BE818805.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 424)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
```

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM2-BN0302-040
700-253-B10&t3=2000-07-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 424.

FEATURES

source
1..424
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"

/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 424;
Best Local Similarity 64.3%; Pred. No. 0.06;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 250 GCCCAGCTTGGCGGCGCCCTTACAAGACTATGAAGTAAAG 209

RESULT 8

BE818807/c
LOCUS
DEFINITION CM2-BN0302-040700-253-d02 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE818807
VERSION BE818807.1 GI:10251041
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 449)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM2-BN0302-040

700-253-d02&t3=2000-07-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 306.

FEATURES

source
1..449
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 449;
Best Local Similarity 64.3%; Pred. No. 0.061;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 240 GCCCAGCTTGGCGGCGCCCTTACAAGACTATGAAGTAAAG 199

RESULT 9

CD672930
LOCUS
DEFINITION fg18a05-y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
fg18a05 5', mRNA sequence.

ACCESSION CD672930
VERSION CD672930.1 GI:32174661
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

1 (bases 1 to 453)
Wistow, G., Bernstein, S.L., Ray, S., Wyatt, M.K., Behal, A.,
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of adult human iris for the NEIBank
Project: steroid-response factors and similarities with retinal
pigment epithelium

JOURNAL Mol. Vis. 8 (4), 185-195 (2002)
PUBMED 12107412
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov

Plate: 18 row: a column: 05
Seq primer: M3RP1 reverse primer (ABI).
Location/Qualifiers
1..453
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fg18a05"
/tissue="iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Iris cDNA (Normalized): fg"
/notes="Organ: Eye; Vector: pCMVSPORT6; A human iris
library (bx) was normalized by self-subtraction. One
portion of double stranded plasmid DNA representing the
library was linearized by NotI. This NotI digested library

FEATURES

source
1..453
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fg18a05"
/tissue="iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Iris cDNA (Normalized): fg"
/notes="Organ: Eye; Vector: pCMVSPORT6; A human iris
library (bx) was normalized by self-subtraction. One
portion of double stranded plasmid DNA representing the
library was linearized by NotI. This NotI digested library

```

1. .458
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lb="EN0302"
/notes="Organ: breast_normal; Vector: puc18; Site1: 5mal; Site2: 5mal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 195,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```



```

RESULT 12
BE818852/c
LOCUS          467 bp      mRNA      linear      EST 21-SEP-2000
DEFINITION    CM2-BN0302-100700-259-a05 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION     BE818852
VERSION       BE818852.1 GI:10251086
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Hominiidae; Homo.
REFERENCE
AUTHORS       1 (bases 1 to 467)
               Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
               Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
               Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
               Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
               O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
               Simpson,A.J.
TITLE         Shotgun sequencing of the human transcriptome with ORF expressed
               sequence tags
JOURNAL       Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED       10737800
COMMENT      Contact: Simpson A.J.G.
               Laboratory of Cancer Genetics
               Ludwig Institute for Cancer Research
               Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
               Brazil
               Tel: +55-11-2704922
               Fax: +55-11-2707001
               Email: asimpson@ludwig.org.br
               This sequence was derived from the FAPESP/LICR Human Cancer Genome
               Project. This entry can be seen in the following URL
               (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=CM2-BN0302-100
               700-259-a05&t3=2000-07-10&t4=1)
               Seq primer: puc 18 forward
               High quality sequence start: 19
               High quality sequence stop: 467.
               Location/Qualifiers
FEATURES
source
1..467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match          90.4%; Score 30; DB 7; Length 467;
Best Local Similarity 64.3%; Pred. No. 0.061;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCNTGGCCNTCCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db 246 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 205
    |||||

RESULT 13
BP253378
LOCUS          571 bp      mRNA      linear      EST 15-SEP-2004
DEFINITION    BP253378 Sugano cDNA library, kidney epithelial cell Homo sapiens
               cDNA clone HRC03720, mRNA sequence.
ACCESSION     BP253378
VERSION       BP253378.1 GI:52135659
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Hominiidae; Homo.
REFERENCE
AUTHORS       1 (bases 1 to 571)
               Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
               Mizushima-Sugano,J., Nakai,K. and Sugano,S.
               Sequence comparison of human and mouse genes reveals a homologous
               block structure in the promoter regions
               Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL       Genomics
PUBMED       15342556
COMMENT      Contact: Yutaka Suzuki
               Department of Virology
               Institute of Medical Science, University of Tokyo
               4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
               Tel: 81-3-5449-5343
               Fax: 81-3-5449-5416
               Email: yusuzuki@ohgc.jp.
               Location/Qualifiers
FEATURES
source
1..571
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC03720"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"
ORIGIN
Query Match          90.4%; Score 30; DB 3; Length 571;
Best Local Similarity 64.3%; Pred. No. 0.064;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCNTGGCCNTCCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db 522 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 563
    |||||

RESULT 14
CB152636
LOCUS          577 bp      mRNA      linear      EST 29-JAN-2003
DEFINITION    K-EST0209868 L12JSHC0s1 Homo sapiens cDNA clone L12JSHC0s1-5-G07
               5', mRNA sequence.
ACCESSION     CB152636
VERSION       CB152636.1 GI:28137590
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Hominiidae; Homo.
REFERENCE
AUTHORS       1 (bases 1 to 577)
               Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
               Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
               Kim,Y.S.
               21C Frontier Korean EST Project 2001
               Unpublished (2002)
JOURNAL       Genomics Research Center
COMMENT      Korea Research Institute of Bioscience & Biotechnology
               52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
               Tel: +82-42-860-4470
               Fax: +82-42-860-4409
               Email: yongsung@mail.kribb.re.kr
               Plate: 5 row: G column: 07
               High quality sequence stop: 577.
               Location/Qualifiers
FEATURES
source
1..577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L12JSHC0s1-5-G07"
/sex="M"

```


KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1. .580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH09144"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/note="thyroid tumor"

ORIGIN
Query Match 90.4%; Score 30; DB 3; Length 580;
Best Local Similarity 64.3%; Pred. No. 0.064;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 538 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAGTAAG 579

RESULT 18
BP256144 582 bp mRNA linear EST 16-SEP-2004
LOCUS
DEFINITION cDNA clone HRC10671, kidney epithelial cell Homo sapiens
ACCESSION BP256144
VERSION BP256144.1 GI:52171374
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN
/clone="HRC10671"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

Query Match 90.4%; Score 30; DB 3; Length 582;
Best Local Similarity 64.3%; Pred. No. 0.064;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 415 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAGTAAG 456

RESULT 19
BP363758 582 bp mRNA linear EST 17-SEP-2004
LOCUS
DEFINITION sapiens cDNA clone TIR02383, mRNA sequence.
ACCESSION BP363758
VERSION BP363758.1 GI:52293963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5343
Fax: 81-3-5449-5416
Email: ysuzuki@hgc.jp.

FEATURES
source
1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TIR02383"
/tissue_type="lung"
/cell_type="fibroblast"
/cell_line="TIG"
/dev_stage="fetal"
/clone_lib="Sugano cDNA library, fetal lung fibroblast TIG"

ORIGIN
Query Match 90.4%; Score 30; DB 3; Length 582;
Best Local Similarity 64.3%; Pred. No. 0.064;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 380 GCCCAGCGTTGGCCGAGCCCTTACAAAGACTATGAGTAAG 421

RESULT 20
BP256147 584 bp mRNA linear EST 16-SEP-2004
LOCUS
DEFINITION cDNA clone HRC10676, mRNA sequence.
ACCESSION BP256147
VERSION BP256147.1 GI:52171377
KEYWORDS EST.

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 584)
AUTHORS    Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
PUBMED     15342556
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Tel: 81-3-5449-5343
            Fax: 81-3-5449-5416
            Email: ysuzuki@hgc.jp.
FEATURES   Location/Qualifiers
            source
            1..584
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="HRC10676"
            /tissue_type="kidney"
            /cell_type="epithelial cell"
            /clone_lib="Sugano CDNA library, kidney epithelial cell"
ORIGIN
Query Match      90.4%; Score 30; DB 3; Length 584;
Best Local Similarity 64.3%; Pred. No. 0.064;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY      1  GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
        |||||
        416  GCCCAGCTTGCCGAGCCCTTACAAAGACTATGAAGTAAAG 457

RESULT 21
LOCUS    CB130950
DEFINITION K-EST0180946 L12JSHC0 Homo sapiens CDNA clone L12JSHC0-5-D03 5',
mRNA sequence.
ACCESSION CB130950
VERSION    CB130950.1 GI:28095412
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 621)
AUTHORS    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.S., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 5 row: D column: 03
            High quality sequence stop: 621.
            Location/Qualifiers
            source
            1..621
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="L12JSHC0-5-D03"

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 630)
AUTHORS    NIH-MGC Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL    COMMENT
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            CDNA library preparation: Ling Hong/Rubin Laboratory
            CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: L12JSHC0 row: o column: 19
            High quality sequence stop: 628.
            Location/Qualifiers
            source
            1..630
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3929706"
            /tissue_type="adenocarcinoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC_9"
            /note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; CDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of

```

California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 630;
Best Local Similarity 64.3%; Pred. No. 0.065;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 66 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 107

RESULT 23

BM843051
LOCUS BM843051
DEFINITION K-ES70120649 S12SNU216 Homo sapiens cDNA clone S12SNU216-77-H12 5',
mRNA sequence.
VERSION BM843051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 637)
Oh, J.H., Yang, J.O., Hahn, Y., Kim, M.R., Byun, S.S., Jeon, Y.J.,
Kim, J.M., Song, K.S., Noh, S.M., Kim, S., Yoo, H.S., Kim, Y.S. and
Kim, N.S.

REFERENCE

Transcriptome analysis of human gastric cancer
Mamm. Genome 16 (12), 942-954 (2005)

COMMENT

Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 77 row: H column: 12
High quality sequence stop: 637.
Location/Qualifiers
1. .637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-77-H12"
/sex="F"
/tissue_type="Lymph node"
/cell_line="Epithelial"
/lab_host="SNU-216"
/clone_lib="S12SNU216"
/notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

FEATURES

source

1. .637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-77-H12"
/sex="F"
/tissue_type="Lymph node"
/cell_line="Epithelial"
/lab_host="SNU-216"
/clone_lib="S12SNU216"
/notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 90.4%; Score 30; DB 3; Length 637;

Best Local Similarity 64.3%; Pred. No. 0.065;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 59 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 100

RESULT 24

BF568108
LOCUS BF568108
DEFINITION 602183908F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4300248 5',
mRNA sequence.
VERSION BF568108
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 699)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLCml158 row: o column: 01

High quality sequence stop: 638.

Location/Qualifiers

1. .699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300248"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

Query Match 90.4%; Score 30; DB 7; Length 699;

Best Local Similarity 64.3%; Pred. No. 0.066;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 321 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 362

FEATURES

source

1. .699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300248"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

ORIGIN

Query Match 90.4%; Score 30; DB 7; Length 699;
Best Local Similarity 64.3%; Pred. No. 0.066;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 321 GCCCAGCGTTGGCGAGCCCTTACAAGACTATGAAGTAAAG 362

RESULT 25

BG251131
LOCUS BG251131
DEFINITION 602364932F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473439 5',
mRNA sequence.
VERSION BG251131
KEYWORDS EST.

BG251131
LOCUS BG251131
DEFINITION 602364932F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473439 5',
mRNA sequence.
VERSION BG251131
KEYWORDS EST.

SOURCE ORGANISM	Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 702)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2570 row: 0 column: 08 High quality sequence stop: 688.
FEATURES source	Location/Qualifiers 1..702 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4473439" /tissue_type="adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 90" /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN	
Query Match	90.4%; Score 30; DB 2; Length 702;
Best Local Similarity	64.3%; Pred. No. 0.066;
Matches	27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY	1 GCNCAYGCGTGGCCNTCCNTAYARGAYTAGTGTTNAAR 42
Db	464 GCCACGCTTGCCGCCGCGCCCTTACAAGACTATGAAGTAAAG 505
RESULT 26	BQ887352
LOCUS	BQ887352
DEFINITION	AGENCOURT_8678064 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381080
ACCESSION	BQ887352
VERSION	BQ887352.1 GI:22279366
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 851)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2570 row: 1 column: 09 High quality sequence stop: 579.
FEATURES source	Location/Qualifiers 1..851 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6721245" /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCap" /lab_host="EMDH10B" /note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Inset size: 1800 bp. J. Vincent, Robert Strausberg, Kristi A. Egland, James J. Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins.
ORIGIN	
Query Match	90.4%; Score 30; DB 3; Length 851;
Best Local Similarity	64.3%; Pred. No. 0.069;
Matches	27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY	1 GCNCAYGCGTGGCCNTCCNTAYARGAYTAGTGTTNAAR 42
Db	182 GCCACGCTTGCCGCCGCGCCCTTACAAGACTATGAAGTAAAG 223
RESULT 27	CA454746
LOCUS	CA454746
DEFINITION	AGENCOURT_10763191 MAPcL Homo sapiens cDNA clone IMAGE:6721245 5', mRNA sequence.
ACCESSION	CA454746
VERSION	CA454746.1 GI:24904781
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 865)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Kristi A. Egland, Ira Pastan cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM14282 row: a column: 21 High quality sequence stop: 686.
FEATURES source	Location/Qualifiers 1..865 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6721245" /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCap" /lab_host="EMDH10B" /note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Inset size: 1800 bp. J. Vincent, Robert Strausberg, Kristi A. Egland, James J. Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins.


```
RESULT 33
BX370558      962 bp      mRNA      linear      EST 27-APR-2004
LOCUS
DEFINITION    BX370558 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK011YF17 5-PRIME, mRNA sequence.
ACCESSION     BX370558
VERSION       BX370558.1 GI:30453927
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 962)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6601.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAG0492C03_CS04627_1&c=6601.f

FEATURES
source
1. .962
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YF17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."

ORIGIN
Query Match 90.4%; Score 30; DB 4; Length 962;
Best Local Similarity 64.3%; Pred. No. 0.071;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 248 GCCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAAAG 289

RESULT 34
BQ067433      1023 bp      mRNA      linear      EST 02-APR-2002
LOCUS
DEFINITION    BQ067433 AGENCOURT_6758944 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754944
5', mRNA sequence.
ACCESSION     BQ067433
VERSION       BQ067433.1 GI:19896479
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1023)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAMI2793 row: C column: 09
High quality sequence stop: 649.

FEATURES
source
1. .1023
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5754944"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 Kb,
insert size range 1-3 Kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 90.4%; Score 30; DB 3; Length 1023;
Best Local Similarity 64.3%; Pred. No. 0.072;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 403 GCCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAAAG 444

RESULT 35
CR597125      1830 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS
DEFINITION    CR597125 full-length cDNA clone CS0DK011YF17 of HeLa cells COT 25-normalized
of Homo sapiens (human).
ACCESSION     CR597125
VERSION       CR597125.1 GI:50477932
KEYWORDS      HTC; CNSiUT_cDNA.
SOURCE        Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1830)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1830)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source
1. .1830
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
```

```

/db_xref="taxon:9606"
/clone="CS0BK011YF17"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="PCWVSPORT_6"

ORIGIN
Query Match          90.4%; Score 30; DB 6; Length 1830;
Best Local Similarity 64.3%; Pred. No. 0.08;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||:||||| |||:|||||:||||:||||:
Db 1118 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 1159

RESULT 36
BI193620
LOCUS
DEFINITION
602946519F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5089900 5',
mRNA sequence.
ACCESSION
BI193620
VERSION
BI193620.1 GI:14648640
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1125)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1851 row: e column: 05
High quality sequence stop: 529.
FEATURES
Location/Qualifiers
1..1125
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5089900"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 42"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match          88.6%; Score 29.4; DB 2; Length 1125;
Best Local Similarity 65.9%; Pred. No. 0.13;
Matches 27; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNA 41
    |||:||||| |||:|||||:||||:||||:
Db 294 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 334

RESULT 37

```

```

CF780868
LOCUS
DEFINITION
AGENCOURT_15939245 NIH_MGC_219 Homo sapiens cDNA clone
IMAGE:30523569 5', mRNA sequence.
ACCESSION
CF780868
VERSION
CF780868.1 GI:37740645
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 916)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhardt, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Martin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM601 row: 1 column: 10
High quality sequence stop: 651.
FEATURES
Location/Qualifiers
1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30523569"
/tissue_type="Pooled Chondrosarcoma Tumor cells"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_219"
/notes="Vector: pYX-Asc; Site 1: EcoRI; Site 2: NotI;
Library is oligo-dT primed and directionally cloned
Denatured RNA was size fractionated on a 1% agarose gel.
First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
0.5-1kb. Adaptors 5'(AATTGGCAGCAGG)3' and 5'd
(CCTGCTGCGG)3'. 3' Linker sequence - GCGGCGCTGAGAGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(AATTACCTCTCAATGAAGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Average insert size 2-3kb.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match          87.3%; Score 29; DB 5; Length 916;
Best Local Similarity 61.9%; Pred. No. 0.19;
Matches 26; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||:||||| |||:|||||:||||:||||:
Db 667 GCCACGCTTGGCCGAGCCCTTACAAGACTATGAAGTAA 708

RESULT 38
BE818831/c
LOCUS
DEFINITION
CM2-BN0302-050700-256-a12 BN0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE818831
VERSION
BE818831.1 GI:10251065
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS 1. (bases 1 to 365)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBLISHED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM2-BN0302-050
700-256-al2&t3=2000-07-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 365.

FEATURES
source
1..365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0302"
/notes="Organ: breast normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 85.5%; Score 28.4; DB 7; Length 365;
Best Local Similarity 61.9%; Pred. No. 0.3;
Matches 26; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GCNCAYGCVTGGCCNTCCNTAYAAAGATAYGARGTNAAR 42
|||:||||| |||:|||||:||||:||||:||||:
Db 267 GCCACGCTGGCGCGCCCTTACAAAGACTATGAAGTAAAG 226

RESULT 39
BO685729
LOCUS BO685729 940 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8344399 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251057
5', mRNA sequence.

ACCESSION BO685729

VERSION BO685729.1 GI:21811045

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 940)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Straubeberg, Ph.D.

Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC2395 row: b column: 18
High quality sequence stop: 538.

FEATURES
source
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6251057"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 85.5%; Score 28.4; DB 3; Length 940;
Best Local Similarity 61.9%; Pred. No. 0.36;
Matches 26; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GCNCAYGCVTGGCCNTCCNTAYAAAGATAYGARGTNAAR 42
|||:||||| |||:|||||:||||:||||:||||:
Db 687 GCCACGCTGGCGCGCCCTTACAAAGACTATGAAGTAAAG 728

RESULT 40
AU140676
LOCUS AU140676 PLACE4 Homo sapiens cDNA clone linear EST 05-AUG-2002
DEFINITION 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
sequence.

ACCESSION AU140676

VERSION AU140676.1 GI:11002197

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 680)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes

TITLE Genome Res. 16 (1), 55-65 (2006)

JOURNAL 16344560

PUBLISHED Contact: Takao Isogai

COMMENT FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and

```

FEATURES
source
    Helix Research Institute
    Location/Qualifiers
    1..680
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="PLACE4000034"
    /tissue_type="placenta"
    /clone_lib="PLACE4"
    /notes="Vector: pME18SFL3"

ORIGIN
    Query Match      82.5%; Score 27.4; DB 1; Length 680;
    Best Local Similarity 59.5%; Pred. No. 0.93;
    Matches 25; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 GNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
    |||||
Db 633 GCCCAGCGTGGCCACCCCTTACAAGACTATGAANGTAA 674

RESULT 41
AA615518
LOCUS
DEFINITION
    v070e02.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
IMAGE:1064474 5', mRNA sequence.
ACCESSION
    AA615518
VERSION
    AA615518.1 GI:2502746
KEYWORDS
    EST.
SOURCE
    Mus musculus (house mouse)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
    1 (bases 1 to 454)
AUTHORS
    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
    Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
    Schellenberg,K., Stepoe,M., Tan,F., Underwood,K., Moore,B.,
    Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
    Waterston,R.
    The WashU-HHMI Mouse EST Project
    Unpublished (1996)
    Contact: Marra M/Mouse EST Project
    WashU-HHMI Mouse EST Project
    Washington University School of MedicineP
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: mouseest@watson.wustl.edu
    This clone is available royalty-free through LML ; contact the
    IMAGE Consortium (info@image.llnl.gov) for further information.
    MGI:586934
    Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
source
    Location/Qualifiers
    1..454
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C3H"
    /db_xref="taxon:10090"
    /clone="IMAGE:1064474"
    /cell_line="C2C12"
    /lab_host="DH108"
    /clone_lib="Barstead mouse myotubes MPLRB5"
    /notes="Vector: pT7T3D-Paci; Site 1: EcoRI; Site 2: NotI;
    1st strand cDNA was primed with a Not I - oligo(dT) primer
    [5',
    TGTTCAGATCTCACTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
    3'] ; double-stranded cDNA was ligated to Eco RI adaptors
    [AATTCGGACCTTG], digested with Not I and cloned into the
    Not I and Eco RI sites of the modified pT7T3 vector.
    Library constructed by Bob Barstead. The C2C12 cell line
    (available from ATCC, catalog # CRL-1772) differentiates
    rapidly, forming contractile myotubes and producing

```

```

ORIGIN
    characteristic muscle proteins. "

Query Match      72.3%; Score 24; DB 1; Length 454;
Best Local Similarity 64.9%; Pred. No. 27;
Matches 24; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 GNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARG 37
    |||||
Db 326 GCCCAGCGCGCCCATGTCCGTACATGACTACGAAG 362

RESULT 42
CA762213
LOCUS
DEFINITION
    CA762213 652 bp mRNA linear EST 27-NOV-2002
    BR060012A20C08.ab1 IIRI clones Oryza sativa (indica cultivar-group)
    cDNA clone BR060012A20C08.ab1 similar to NP.187283.1 [NM.111507]
    putative 50S ribosomal protein L33 [Arabidopsis, mRNA sequence.
CA762213
ACCESSION
    CA762213.1 GI:25806252
VERSION
    CA762213.1
KEYWORDS
    EST.
SOURCE
    Oryza sativa (indica cultivar-group)
ORGANISM
    Oryza sativa (indica cultivar-group)
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
    clade; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
    1 (bases 1 to 652)
AUTHORS
    Bennett,J., Arumugam,K., Lafitte,R., Wen,J. and Bruskiewich,R.
    Rice Microarray
    Unpublished (2002)
    Contact: Mark Fredricksen
    Department of Plant Biology
    University of Illinois
    1201 W. Gregory Dr., Urbana, IL 61801, USA
    Tel: 2172655473
    Email: bohnerl@life.uiuc.edu
    International Rice Information System (IRIS)
    http://www.iris.irri.org).
    Location/Qualifiers
    1..652
    /organism="Oryza sativa (indica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="IR64"
    /db_xref="taxon:39946"
    /clone="BR060012A20C08.ab1"
    /tissue_type="Panicles"
    /dev_stage="Flowering"
    /clone_lib="IRRI Clones"
    /note="Vector: pBluescript II SK+; Water stress was
    applied by not watering for 4 consecutive days. Panicles
    were collected from control (well watered) and stressed
    plants at 2 days before heading, heading, 50% flowering
    and 4 days after 50% flowering."

FEATURES
source
    Location/Qualifiers
    1..652
    /organism="Oryza sativa (indica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="IR64"
    /db_xref="taxon:39946"
    /clone="BR060012A20C08.ab1"
    /tissue_type="Panicles"
    /dev_stage="Flowering"
    /clone_lib="IRRI Clones"
    /note="Vector: pBluescript II SK+; Water stress was
    applied by not watering for 4 consecutive days. Panicles
    were collected from control (well watered) and stressed
    plants at 2 days before heading, heading, 50% flowering
    and 4 days after 50% flowering."

ORIGIN
    Query Match      71.1%; Score 23.6; DB 4; Length 652;
    Best Local Similarity 59.0%; Pred. No. 44;
    Matches 23; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
    |||||
Db 426 CACGCTTTCCTTCCCATCTATGTTATGAGCCAA 464

RESULT 43
CNS02R3E
LOCUS
DEFINITION
    Tetraodon nigroviridis genome survey sequence T7 end of clone
    159K05 of library G from Tetraodon nigroviridis, genomic survey
    sequence.
ACCESSION
    AL210083
VERSION
    AL210083.1 GI:7868902
KEYWORDS
    GSS; genome survey sequence.

```


Db 1058 CCCACGCTGGCCGTCTCCCTGTTAAGACCAGGCGGACAA 1097

RESULT 48	DT989670	1439 bp	mrna	linear	EST 22-SEP-2005
LOCUS	CUJ245-B03.y1d-s	SHGC-CLJ	Gasterosteus aculeatus	cdna clone	
DEFINITION	CUJ245-B03 5',	mrna sequence.			
ACCESSION	DT989670				
VERSION	DT989670.1	GI:76113480			
KEYWORDS	EST.				
SOURCE	Gasterosteus aculeatus	(three spined stickleback)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.				
REFERENCE	1	(bases 1 to 1439)			
AUTHORS	Kingley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M., Schmutz, J., and Myers, R.M.				
TITLE	Expressed sequence tags from Gasterosteus aculeatus				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Grimwood, Jane				
	Stanford Human Genome Center				
	Stanford University School of Medicine				
	975 S California Ave, Palo Alto, CA 94304, USA				
	Tel: 650 320 5917				
	Fax: 650 320 5801				
	Email: jane@shgc.stanford.edu				
	Plate: 245				
	High quality sequence stop: 592.				

```

FEATURES
source
    Location/Qualifiers
        1..1439
            /organism="Gasterosteus aculeatus"
            /mol_type="mRNA"
            /strain="Bitruufjordur marine sticklebacks, Iceland"
            /db_xref="taxon:69293"
            /clone="CLJ245-B03"
            /sex="mixed male and female"
            /tissue_type="whole larva"
            /dev_stage="21 day old larvae collected at Swarup Stage 30
(J. Embryol. Exp. Morphol 6: 373-383.1958)"
            /lab_host="DH10B (T1 phage resistant)"
            /clone_lib="SHGC-CLJ"
            /note="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed by
with an 54 bp linker primer containing an oligodt sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTTCAGATCGGAGCGCCGCC(T)25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction.fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria. Clones available
from Open Biosystems:
http://www.openbiosystems.com/stickleback"

```

ORIGIN
http://www.openprosystems.com/stickiepack"

Query Match	67.5%	Score 22.4;	DB 10;	Length 1439;
Best Local Similarity	57.5%;	Pred. No. 1.8e+02;		
Matches 23; Conservative	5;	Mismatches 12;	Indels 0;	Gaps 0;
QY	2	CNCAYGNTGGCCNCCNTAYAARGAYTAYGARGTNAA	41	
		: :		
db	133	CTCCGCATGGGGCTTCGCCGTATAAGAAATGATCCTGTGA	172	
		: :		

RESULT 49	
BE006062	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

REFERENCE
AUTHORS

TITLE

JOURNAL
PUBMED
COMMENT

FEATURES
source

Strong

ORIGIN

Query Match	66.9%	Score 22.2;	DB 7;	Length 583;
Best Local Similarity	65.6%	Pred. No. 1.8e+02;		
Matches 21: Conservative	4;	Mismatches 7;	Indels 0;	Gaps 0;

Qy 1 GCNCAYGCNTGGCCNCTCNCNTTAYAARGAYTA 32
|||:||| ||| ||| :||| :|||
Db 552 GCCCACGCTTGGCCGAGGCCCTTACAAAGACTA 583

RESULT 50
AQ692009/c

LOCUS	AQ692009	560 bp	DNA	linear	GSS 06-JUL-1999
DEFINITION	HS 5402_B1_H12_T7A_RPCT-11 Human Male BAC Library Homo sapiens genomic clone Plate=978 Col=23 Row=P, genomic survey sequence.				
ACCESSION	AQ692009				
VERSION	AQ692009.1	GI:5382257			
KEYWORDS	GSS.				

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 560)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc.Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 1049764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 978 row: P column: 23
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 560.
FEATURES
 source
 1..560
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=978 Col=23 Row=P"
 /sex="male"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"
ORIGIN
 Query Match 66.3%; Score 22; DB 11; Length 560;
 Best Local Similarity 52.4%; Pred. No. 2.2e+02;
 Matches 22; Conservative 7; Mismatches 13; Indels 0; Gaps 0;
 QY 1 GCNCGTCNTGCCNTCCNTAYARGAYTAYGAGTNAAR 42
 |||||
 Db 198 GCCCATTCATGGACATCTCCACACAAAGTCTACGGACTTTAA 157
 |||||
RESULT 51
 CL815091 805 bp DNA linear GSS 09-AUG-2004
LOCUS OR_Cba0031004.r OR_Cba Oryza rufipogon genomic clone OR_Cba0031004
DEFINITION 3'-genomic survey sequence.
ACCESSION CL815091
VERSION CL815091.1 GI:51058756
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 805)
AUTHORS Kim,H., Yu,Y., Wisseotaki,M., Yost,D., Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE ONAP project
JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 PCR Primers
 FORWARD: TAA TAC GAC TCA CTA TAG GG
 BACKWARD: CAC TCA TTA GGC ACC CCA
 Plate: 0031 row: O column: 04
 Seq primer: CAC TCA TTA GGC ACC CCA
 Class: BAC ends.
FEATURES
 Location/Qualifiers
 1..805
 /organism="Oryza rufipogon"
 /mol_type="genomic DNA"
 /db_xref="taxon:4529"
 /clone="OR_CBa0031004"
 /tissue_type="young leaves"
 /dev_stage="2 week old seedlings"
 /lab_host="DH10B T1 phage resistant"
 /clone_lib="OR_CBa"
 /notes="Vector: pAGTBA1; Site 1: HindIII; Site 2: HindIII; drk treated 36 hrs before harvest"
ORIGIN
 Query Match 66.3%; Score 22; DB 13; Length 805;
 Best Local Similarity 56.4%; Pred. No. 2.4e+02;
 Matches 22; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 QY 4 CAYGNTGCCNTCCNTAYARGAYTAYGAGTNAAR 42
 |||||
 Db 84 CACGCTTTCCTTCCCTCATCTAGTGTATGAGCCAA 122
 |||||
RESULT 52
 AG837338/c 842 bp DNA linear GSS 03-NOV-2004
LOCUS Oryza sativa (indica cultivar-group) genomic DNA, BAC end sequence,
 DEFINITION BAC clone:K0038B07_R, genomic survey sequence.
ACCESSION AG837338
VERSION AG837338.1 GI:55303573
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Katagiri,S., Wu,J., Ito,Y., Karasawa,W., Shibata,M., Kanamori,H., Katayose,Y., Namiki,N., Matsumoto,T. and Sasaki,T.
TITLE End Sequencing and Chromosomal in silico Mapping of BAC Clones
 Derived from an indica Rice Cultivar, Kasalath
JOURNAL Breeding Science 54, 273-279 (2004)
REFERENCE 2 (bases 1 to 842)
AUTHORS Sasaki,T., Matsumoto,T. and Wu,J.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2004) Takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba Ibaraki, 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT The orientation of the sequence is from SP6 side of the BAC clone.
FEATURES
 Location/Qualifiers
 1..842
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Kasalath"
 /db_xref="taxon:39946"
 /clone="K0038B07_R"
ORIGIN

```

Query Match      66.3%; Score 22; DB 14; Length 842;
Best Local Similarity 56.4%; Pred. No. 2.4e+02;
Matches 22; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 4 CAYCNGTGGCCNTCNCNTAYAARGATAYAGTGAAR 42
Db 126 CAGCCTTGGCTTCCCATCTAGTGGTTATGAAGCCAAA 88

RESULT 53
LOCUS CV968179/c 880 bp mRNA linear EST 25-JAN-2005
DEFINITION PC057H02 infected tomato, center of lesion 3 dpi Phytophthora
infectans cDNA, mRNA sequence.
ACCESSION CV968179
VERSION 1
KEYWORDS EST.
SOURCE Phytophthora infestans (potato late blight agent)
ORGANISM Phytophthora infestans
REFERENCE 1 (bases 1 to 880)
AUTHORS Randall.T., Dwyer.R.A., Huitema.E., Beyer.K., Cvitanich.C.,
Kelkar.H., Fong.A.M., Gates.K., Roberts.S., Yatzkan.E., Gaffney.T.,
Law.M., Testa.A., Torco-Alalibo.A., Zhang.M., Zheng.L., Mueller.E.,
Windass.J., Binder.A., Birch.P.R.J., Gisi.U., Govers.F., Gow.N.A.,
Mauch.F., van West.P., Waugh.M.E., Yu.J., Boller.T., Kamoun.S.,
Lam.S.T. and Judelson, H.S.
TITLE Large-scale gene discovery in the oomycete Phytophthora infestans
reveals likely components of phytopathogenicity shared with true
fungi
JOURNAL Mol. Plant-Microbe Interact. 18 (3), 229-243 (2005)
PUBMED 15782637
COMMENT Contact: Judelson HS
Department of Plant Pathology
University of California
Webster Hall, Riverside, CA 92521, USA
Tel: 909 787 4199
Fax: 909 787 4294
Email: howard.judelson@ucr.edu
may be of host plant origin.
FEATURES
source Location/Qualifiers
1..880
/mol_type="mRNA"
/strain="88069"
/db_xref="taxon:4787"
/sex="Al"
/clone_lib="infected tomato, center of lesion 3 dpi"
/notes="Vector: pSPORT1"
ORIGIN
Query Match      66.3%; Score 22; DB 8; Length 880;
Best Local Similarity 52.4%; Pred. No. 2.4e+02;
Matches 22; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GCNCAYGCTGGCCNTCNCNTAYAARGATAYAGTGAAR 42
Db 684 GCCAATAGTGGCCCTCACCTATCAAGTTTGGGGTCGAG 643

RESULT 54
LOCUS CD190259/c 315 bp mRNA linear EST 14-SEP-2003
DEFINITION MS1-0063U-A372-E08-U.G MS1-0063 Schistosoma mansoni cDNA clone
MS1-0063U-A372-E08.G, mRNA sequence.
ACCESSION CD190259
VERSION 1
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 315)

Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
Sa,R.G., Stukart,G., Soares,M.B., Gargioni,C., Kawano,T.,
Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
Transcriptome analysis of the acelomate human parasite Schistosoma
mansoni
Nat. Genet. 35 (2), 148-157 (2003)
LOCUS 12973350
COMMENT Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquímica
Instituto de Química - Universidade de São Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,
Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjo@iq.usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST
Genome Project. All sequences in the project were assembled and
annotated. This entry and all the assembled sequences can be seen
in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: MS1-0063U-A372 row: 8 column: E.
FEATURES
source Location/Qualifiers
1..315
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MS1-0063U-A372-E08.G"
/sex="mixed pool"
/dev_stage="schistosomulum"
/lab_host="in vitro culture"
/clone_lib="MS1-0063"
/notes="Vector: pGEM T-easy"
ORIGIN
Query Match      65.7%; Score 21.8; DB 5; Length 315;
Best Local Similarity 60.5%; Pred. No. 2.4e+02;
Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCNCAYGCTGGCCNTCNCNTAYAARGATAYAGTGAAR 38
Db 58 GCATGATGTCGCCAGGACCATCAATGATGATGATGT 21

RESULT 55
LOCUS DU888204 936 bp DNA linear GSS 03-JAN-2006
DEFINITION 389797 Tomato HindIII BAC Library Lycopersicon esculentum genomic
clone LE HBa0211C12 3, genomic survey sequence.
ACCESSION DU888204
VERSION 1
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 936)
AUTHORS Mueller,L.A., Buels,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
Van Eck,J. and Stack,S.
TITLE BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL Unpublished (2005)
COMMENT Other GSSs: 389798
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683

```

[illegible]

University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 Plate: 0166 row: J column: 17
 Class: BAC ends.

FEATURES

source

Location/Qualifiers

1. 734
 /organism="Oryza granulata"
 /mol_type="genomic DNA"
 /db_xref="taxon:110450"
 /clone="OG_Aba0166J17"
 /tissue types="young leaves"
 /lab host="DH10B T1 phage resistant"
 /clone lib="OG_ABa"
 /note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 65.1%; Score 21.6; DB 14; Length 734;
 Best Local Similarity 56.8%; Pred. No. 3.5e+02;
 Matches 21; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
 Qy 2 CNGCGTGGCCNTCNCNTAYAARGATYAGT 38
 Db 365 CTCTGTCATGCCATCACCAATATGTCATCAAGT 401

RESULT 59

LOCUS

CZ983576 859 bp DNA linear GSS 11-AUG-2005
 197310 Tomato MboI BAC Library Lycopersicon esculentum genomic
 clone SL_MboI0086A08 5, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (Solanum lycopersicum)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

AUTHORS

1 (bases 1 to 859)
 Mueller,L.A., Buel,S.R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
 Van Eck,J. and Stack,S.
 BAC end sequencing from three Solanum lycopersicon libraries
 Unpublished (2005)

TITLE

JOURNAL

COMMENT

Contact: Lukas Mueller
 Tankley Lab, Dept. of Plant Breeding
 Cornell University
 251 Emerson Hall, Ithaca, NY 14853, USA
 Tel: 607-255-6557
 Fax: 607-255-6683
 Email: sgn-feedback@sgn.cornell.edu
 Plate: 86 row: A column: 8
 Seq primer: T7

Class: BAC ends

High quality sequence start: 59

High quality sequence stop: 316.

FEATURES

source

1. 859
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="SL_MboI0086A08"
 /lab host="E. coli"
 /clone lib="Tomato MboI BAC Library"
 /note="Vector: pBelOBAC11; Site_1: MboI"

ORIGIN

Query Match 65.1%; Score 21.6; DB 13; Length 859;
 Best Local Similarity 61.8%; Pred. No. 3.6e+02;

Matches

21; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy

5 AYGCTGGCCNTCNCNTAYAARGATYAGT 38

Db

354 AGCATGGCCACCCCATCCAAAGATTCTGAAGT 387

RESULT 60

LOCUS

DR753481/c 978 bp mRNA linear EST 21-JUL-2005
 CCRAP4A04 Coprinus cinereus rapamycin-treated mycelia cDNAs
 Coprinopsis cinerea cDNA, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

DR753481.1 GI:71045923
 EST.
 Coprinopsis cinerea (Coprinus cinereus)
 Coprinopsis cinerea
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 Agaricales; Psathyrellaceae; Coprinopsis.

REFERENCE

AUTHORS

1 (bases 1 to 978)
 Carlson,M.D., Barr,C., Murphy,B., Gathman,A.C., Lilly,W.W. and
 Pukkila,P.J.

TITLE

JOURNAL

COMMENT

Expressed sequence tags from Coprinus cinereus (Coprinopsis
 cinerea) cDNAs, summer 2004
 Unpublished (2004)
 Contact: Gathman AC
 Biology Department
 Southeast Missouri State University
 1 University Plaza, Cape Girardeau, MO 63701, USA
 Tel: 5736512361
 Fax: 573 651 2382
 Email: agathman@semo.edu.

FEATURES

source

Location/Qualifiers

1. 978
 /organism="Coprinopsis cinerea"
 /mol_type="mRNA"
 /strain="Okayama7#130"
 /db_xref="taxon:5346"
 /day stage="vegetative monokaryotic mycelium"
 /lab host="E. coli XL10-Gold"
 /clone lib="Coprinus cinereus rapamycin-treated mycelia
 cDNAs"
 /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN

Query Match 65.1%; Score 21.6; DB 10; Length 978;

Best Local Similarity 53.8%; Pred. No. 3.7e+02;

Matches 21; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy

4 CAYGCTGGCCNTCNCNTAYAARGATYAGTNAAR 42

Db

371 CACGTTTGCCCTTCCTCTTACAGGCCCATGGAGTTGAA 333

RESULT 61

LOCUS

CK084463/c 583 bp mRNA linear EST 05-FEB-2005
 79195rsicen.2839.y1 Oryza sativa cv. LYP9 booting whole plant cDNA
 library Oryza sativa (indica cultivar-group) cDNA 5', mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

CK084463.1 GI:58695776
 EST.
 Oryza sativa (indica cultivar-group)
 Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoideae; Oryzae; Oryza.

REFERENCE

AUTHORS

1 (bases 1 to 583)
 Yu,J., Wang,J., Lin,W., Li,S., Li,H., Zhou,J., Ni,P., Dong,W.,
 Hu,S., Zeng,C., Zhang,J., Zhang,Y., Li,R., Xu,Z., Li,S., Li,X.,
 Zheng,H., Cong,L., Lin,L., Yin,J., Geng,J., Li,G., Shi,J., Liu,J.,

Lv,H., Li,J., Wang,J., Deng,Y., Ran,L., Shi,X., Wang,X., Wu,Q.,
Li,C., Ren,X., Wang,J., Wang,X., Li,D., Liu,D., Zhang,X., Ji,Z.,
Zhao,W., Sun,Y., Zhang,Z., Bao,J., Han,Y., Dong,L., Ji,J., Chen,P.,
Wu,S. and Liu,J.

The Genomes of *Oryza sativa*: A History of Duplications

PLoS Biol. 3 (2), e38 (2005)

TITLE

JOURNAL

PUBMED

COMMENT

Contact: Yan Zhou

Bioinformatics Department

Hangzhou Genomics Institute

No.51 ZhiJiang Road, Hangzhou 310008, China

Tel: 86-571-58805886

Fax: 86-571-56805884

Email: zhouyan@genomics.org.cn

Seq primer: M13 Forward

High quality sequence stop: 583

POLYA=No.

Location/Qualifiers

1..583

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="mRNA"

/cultivar="LYP9"

/db_xref="taxon:39946"

/issue_type="whole plant"

/dev_stage="booting"

/clone_lib="Oryza sativa cv. LYP9 booting whole plant cDNA library"

ORIGIN

Query Match

Best Local Similarity 64.5%; Score 21.4; DB 5; Length 583;

Matches 22; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY

2 CMCAYCCTGGCCTCCNTCCNTAYAARGAYTAGTGTAAR 42

Db

100 CGCTCGCCTCGCGCTGCCATCTAAATATTCAGAGGTTTCAG 60

RESULT 62

BZ990137

LOCUS

DEFINITION BZ990137 756 bp DNA linear GSS 25-MAR-2003

PUGIF84TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa390M24,

genomic survey sequence.

ACCESSION BZ990137

VERSION BZ990137.1 GI:29230635

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 756)

REFERENCE

AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and

Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Other GSSs: PUGIF84TD

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..756

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTa390M24"

/clone_lib="ZM_0.6_1.0_KB"

/note=Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

CoT selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity 64.5%; Score 21.4; DB 12; Length 756;

Matches 19; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY

12 GCCTCNCNTAYAARGAYTAGTGTAAR 42

Db

645 GCCTTCTCCATATATAAATTATGAGTTAA 675

RESULT 63

CC716597/c

LOCUS

DEFINITION CC716597 774 bp DNA linear GSS 19-JUN-2003

OGUIZ72TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0452K24,

genomic survey sequence.

ACCESSION CC716597

VERSION CC716597.1 GI:32121373

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 774)

REFERENCE

AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGUIZ72TV

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: methylation filtered.

Location/Qualifiers

1..774

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0452K24"

/clone_lib="ZM_0.7_1.5_KB"

/note=Vector: pBCSK; Site_1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match

Best Local Similarity 64.5%; Score 21.4; DB 12; Length 774;

Matches 19; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY

12 GCCTCNCNTAYAARGAYTAGTGTAAR 42

Db

265 GCCTTCTCCATATATAAATTATGAGTTAA 235

RESULT 64

CB673669/c

LOCUS

DEFINITION CB673669 792 bp mRNA linear EST 09-APR-2003

OSJNEe08E24.f OSJNEe *Oryza sativa* (japonica cultivar-group) cDNA

clone OSJNEe08E24 5', mRNA sequence.

ACCESSION CB673669

VERSION CB673669.1 GI:29677394

KEYWORDS EST.

SOURCE *Oryza sativa* (japonica cultivar-group)

ORGANISM *Oryza sativa* (japonica cultivar-group)


```

LOCUS      AQ461894               544 bp      DNA          linear      GSS 23-APR-1999
DEFINITION HS_5055_A2_B07_T7A RPCI-11 Human Male BAC Library Homo sapiens
            Genomic clone Plate=631 Col=14 Row=C, genomic survey sequence.
ACCESSION   AQ461894
VERSION     AQ461894.1 GI:4634664
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 544)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Kellier,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequencing-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED     10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieterdejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 631 row: C column: 14
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 544.
FEATURES             source
            1..544
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="plate=631 Col=14 Row=C"
                /sex="male"
                /clone_lib="RPCI-11 Human Male BAC Library"
                /notes="vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
                Male blood DNA was isolated from one randomly chosen donor
                and partially digested with a combination of EcoRI and
                EcoRI Methylase. Size selected DNA was cloned into the
                pBACE3.6 vector at EcoRI sites"
ORIGIN
Query Match      63.3%; Score 21; DB 11; Length 544;
Best Local Similarity 59.5%; Pred. No. 6.1e+02;
Matches 22; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY      6 YGNTGGCCNTCCNTAYARGAYTAGGTAAR 42
      : ||| ||| ||| ||| ||| ||| ||| |||
Db      133 TGTGTGCCCCCTCTCCNTGAAAGACTACGACNTAAAG 97

RESULT 72
DX336552
LOCUS      DX336552               573 bp      DNA          linear      GSS 20-JAN-2006
DEFINITION OR_Aba0262A04.r OR_Aba Oryza ridleyi genomic clone OR_Aba0262A04
            3', genomic survey sequence.
ACCESSION   DX336552
VERSION     DX336552.1 GI:85613338
KEYWORDS    GSS.
SOURCE      Oryza ridleyi
ORGANISM    Oryza ridleyi
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.

```

```

REFERENCE   1 (bases 1 to 573)
AUTHORS     Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
TITLE       OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
            Unpublished (2005)
JOURNAL
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            Plate: 0262 row: A column: 04
            Class: BAC ends.
FEATURES             Location/Qualifiers
            1..573
                /organism="Oryza ridleyi"
                /mol_type="genomic DNA"
                /db_xref="taxon:83308"
                /clone="OR_Aba0262A04"
                /tissue_type="leaves"
                /lab_host="DH10B"
                /clone_lib="OR_Aba"
                /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
Query Match      63.3%; Score 21; DB 14; Length 573;
Best Local Similarity 61.8%; Pred. No. 6.1e+02;
Matches 21; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      2 CMCAYCCTGGCCNTCCNTCCNTAYARGAYTAGGA 35
      ||| ||| ||| ||| ||| ||| ||| |||
Db      377 CACACATGCACATCTCCATATAGGAGGTATGA 410

RESULT 73
AQ727367/c
LOCUS      AQ727367               583 bp      DNA          linear      GSS 14-JUL-1999
DEFINITION HS_5454_A1_D08 SP6E RPCI-11 Human Male BAC Library Homo sapiens
            Genomic clone Plate=1030 Col=15 Row=G, genomic survey sequence.
ACCESSION   AQ727367
VERSION     AQ727367.1 GI:5487036
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 583)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Kellier,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequencing-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED     10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieterdejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1030 row: G column: 15
            Seq primer: SP6
            Class: BAC ends

```

High quality sequence stop: 583.
 Location/Qualifiers
 1. 583
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=1030 Col=15 Row=G"
 /sex="male"
 /clone.lib="RPCI-11 Human Male BAC Library"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"

ORIGIN
 Query Match 63.3%; Score 21; DB 11; Length 583;
 Best Local Similarity 51.2%; Pred. No. 6.1e+02;
 Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;
 QY 2 CNCAYGCTGGCCNTCCNTAATGAAGATATGATGAGTGA 400
 Db 440 CGATGACGGCTCCCAAGATATGATGAGTGA 400

RESULT 74
 CW064489
 LOCUS
 DEFINITION 602 bp DNA linear GSS 28-OCT-2004
 104310_10522599 114_30141 Sorghum methylation filtered library
 (LibID: 104) Sorghum bicolor genomic clone 10522599, genomic survey
 sequence.
 CW064489
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Sorghum bicolor (sorghum)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 602)
 Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
 Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
 McMenamy, J., Smith, M., Holsman, H., Roe, B.A., Wiley, G., Korf, I.F.,
 Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddloh, J.A. and
 Martensen, R.A.
 Sorghum genome sequencing by methylation filtration
 PLOS Biol. 3 (1), e13 (2005)
 15660154
 Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: 310 row: m column: 13
 Seq primer: M13/pUC Forward
 Class: methylation filtered
 High quality sequence stop: 602.
 Location/Qualifiers
 1. 602
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /cultivar="ATx623"
 /db_xref="taxon:4558"
 /clone="10522599"
 /clone.lib="Sorghum methylation filtered library (LibID:
 104)"
 /note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
 prepared from purified nuclei was randomly sheared,
 end-repaired, size fractionated to enrich for the 0.5 to 5
 kb fraction, ligated into HincII-digested pBCSK(-) vector
 and electroporated into E. coli cells. This is a
 methylation filtered library."

FEATURES
 source
 1. 602
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /cultivar="ATx623"
 /db_xref="taxon:4558"
 /clone="10522599"
 /clone.lib="Sorghum methylation filtered library (LibID:
 104)"
 /note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
 prepared from purified nuclei was randomly sheared,
 end-repaired, size fractionated to enrich for the 0.5 to 5
 kb fraction, ligated into HincII-digested pBCSK(-) vector
 and electroporated into E. coli cells. This is a
 methylation filtered library."

ORIGIN
 Query Match 63.3%; Score 21; DB 13; Length 602;
 Best Local Similarity 58.3%; Pred. No. 6.2e+02;
 Matches 21; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 2 CNCAYGCTGGCCNTCCNTAATGAAGATATGAGTGA 37
 Db 277 CGCCGGTAGCGCTGCCGACAGGACGACGACG 312

RESULT 75
 CN249531
 LOCUS
 DEFINITION 615 bp mRNA linear EST 09-APR-2004
 EST015448 Mycelium and yeast cells from Paracoccidioides
 brasiliensis Paracoccidioides brasiliensis cDNA, mRNA sequence.
 CN249531
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Paracoccidioides brasiliensis
 Paracoccidioides brasiliensis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; mitosporic Onygenales; Paracoccidioides.
 1 (bases 1 to 615)
 Felipe, M.S., Andrade, R.V., Arraes, F.B., Nicola, A.M., Matanhao, A.Q.,
 Torres, F.A., Silva-Pereira, I., Pocas-Fonseca, M.J., Campos, E.G.,
 Moraes, L.M., Andrade, P.A., Tavares, A.H., Silva, S.S., Kyaw, C.M.,
 Souza, D.P., Network, P., Pereira, M., Jesuino, R.S., Andrade, E.V.,
 Parente, J.A., Oliveira, G.S., Barbosa, M.S., Martins, N.F.,
 Pachin, A.L., Cardoso, R.S., Passos, G.A., Almeida, N.F., Walter, M.E.,
 Soares, C.M., Carvalho, M.J. and Brigido, M.M.
 Transcriptional Profiles of the Human Pathogenic Fungus
 Paracoccidioides brasiliensis in Mycelium and Yeast Cells
 J. Biol. Chem. 280 (26), 24706-24714 (2005)
 15849188
 Contact: Felipe MSS
 Laboratory of Molecular Biology
 Institute of Biology - University of Brasilia
 Campus Universitario, Asa Norte, Brasilia, DF 70910-900, BRA
 Tel: 55 61 307 2423
 Fax: 55 61 349 8411
 Email: msueli@unb.br
 Seq primer: T7 Sequencing primer.
 Location/Qualifiers
 1. 615
 /organism="Paracoccidioides brasiliensis"
 /mol_type="mRNA"
 /strain="Pb01"
 /db_xref="taxon:121759"
 /clone.lib="Mycelium and yeast cells from Paracoccidioides
 brasiliensis"
 /note="Pb Lambda Zap Express Library"

ORIGIN
 Query Match 63.3%; Score 21; DB 8; Length 615;
 Best Local Similarity 51.2%; Pred. No. 6.2e+02;
 Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;
 QY 2 CNCAYGCTGGCCNTCCNTAATGAAGATATGAGTGA 42
 Db 492 CTCATTCTCTGGCGCTCCAAATAAATACTACGCACTCAA 532

RESULT 76
 CZ770405/c
 LOCUS
 DEFINITION 659 bp DNA linear GSS 26-JUL-2005
 OC_Ba0127H15.r OC_Ba Oryza coarctata genomic clone OC_Ba0127H15
 3'-genomic survey sequence.
 CZ770405
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza coarctata (Porteresia coarctata)
 Oryza coarctata

ORIGIN

Query Match 63.3%; Score 21; DB 13; Length 687;
 Best Local Similarity 58.3%; Pred. No. 6.3e+02;
 Matches 21; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTGGCCNTCCNCTAYAAAGATYAYGARG 37
 Db 149 CGCGCGGTAGCGTCGCGCCGACAGGACGACGACG 114

RESULT 79

DX201680 751 bp DNA linear GSS 19-JAN-2006
 LOCUS OR_Aba0071B03.f OR_Aba Oryza ridleyi genomic clone OR_Aba0071B03
 DEFINITION 5', genomic survey sequence.

ACCESSION DX201680
 VERSION DX201680.1 GI:85451106
 KEYWORDS GSS.

SOURCE

ORGANISM Oryza ridleyi
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS Kim.H., Collura,K., Missotski,M., Byrne,M., Stum,D., Smart,D.,
 Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
 Wing,R.

TITLE ONAP (Oryza Map Alignment Project) - Arizona Genomics Institute

JOURNAL

COMMENT Unpublished (2005)

Contact: Rod A. Wing

Arizona Genomics Institute

University of Arizona

Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: rwing@genome.arizona.edu

Plate: 0071 row: B column: 03

Class: BAC ends.

FEATURES

source Location/Qualifiers
 1..751
 /organism="Oryza ridleyi"
 /mol_type="genomic DNA"
 /db_xref="taxon:83308"
 /clone="OR_Aba0071B03"
 /tissue_type="leaves"
 /lab_host="DH10B"
 /clone_lib="OR_Aba"
 /notes="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 63.3%; Score 21; DB 14; Length 751;
 Best Local Similarity 55.3%; Pred. No. 6.5e+02;
 Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 4 CAYGCGTGGCCNTCCNCTAYAAAGATYAYGARGTNA 41
 Db 365 CTTGATGGCCATCACCCNATCCGATTACCAAGTGCA 402

RESULT 80

DW547742/c 794 bp mRNA linear EST 17-JAN-2006
 LOCUS EST_ssal_rgb2_12161_rgb2 Salmo salar cDNA clone
 DEFINITION ssal_rgb2_520_196_rev 5', mRNA sequence.

ACCESSION DW547742

VERSION DW547742.1 GI:85019086

KEYWORDS EST.

SOURCE

ORGANISM Salmo salar (Atlantic salmon)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.

REFERENCE

1 (bases 1 to 794)
 Koop,B.F., Davidson,W.S. and cGRASP Consortium.

Expressed Sequence Tags from Atlantic Salmon Unpublished (2006)

Unpublished (2006)

COMMENT

Contact: Koop BF

Centre for Biomedical Research

University of Victoria

PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada

Tel: 250 472 4067

Fax: 250 472 4075

Email: bkoop@uvic.ca

Sequenced by: Genome Sciences Centre, BC Cancer Agency Rob Holt, M.
 Marra. Bioinformatics: Centre for Biomedical Research, University
 of Victoria Jong Leong, BF Koop.

Insert Length: 794 Std Error: 0.00

Plate: 520

Seq primer: M13 Reverse

High quality sequence stop: 794.

FEATURES

source Location/Qualifiers
 1..794
 /organism="Salmo salar"
 /mol_type="mRNA"
 /strain="McConnell"
 /db_xref="taxon:8030"
 /clone="ssal_rgb2_520_196_rev"
 /tissue_type="mixed tissue"
 /clone_lib="rgb2"
 /note="Organ: brain, kidney, spleen; Vector: pCMVaport6;
 ssalrgb2 mixed tissue Salmo salar cDNA; Tissue
 contributors: Robert Devlin (DFO, Vancouver, B.C.)"

ORIGIN

Query Match 63.3%; Score 21; DB 10; Length 794;
 Best Local Similarity 55.3%; Pred. No. 6.5e+02;
 Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCCNCTAYAAAGATYAYGARGT 38
 Db 650 GCACACTCTGGCACTCTCCACACAGCCCTCATGAGGT 613

RESULT 81

CW825776

LOCUS CW825776

DEFINITION Ynhw9057 HW-YURAC Bos taurus genomic clone

HW-YURAC2-107-03-E09-T7.ab1 5', genomic survey sequence.

ACCESSION CW825776

VERSION CW825776.1 GI:74045687

KEYWORDS GSS.

SOURCE

ORGANISM Bos taurus (cattle)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Ruminantia; Laurasiatheria; Cetartiodactyla; Ruminantia;

Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 813)

Genomic sequences from Korean Cattle (Hanwoo) blood

Unpublished (2003)

Contact: Inho Choi

Molecular Biology

Yeungnam University

214-1, Dae-dong Gyeongsan, Korea 712-749

Tel: 82 53 810 2933

Fax: 82 53 816 3637

Email: inhochoi@yumail.ac.kr

Seq primer: T7 Forward

Class: BAC ends.

FEATURES

source Location/Qualifiers
 1..813
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="Korean Cattle (Hanwoo)"
 /db_xref="taxon:9913"

/clone="HW-YUBAC2-107-03-E09-T7.ab1"
 /sex="Male"
 /tissue_type="Blood"
 /cell_type="Leucocyte"
 /dev_stage="Adult"
 /clone_lib="HW-YUBAC"
 /note="Vector: pIndigoBAC-5"

ORIGIN

Query Match 63.3%; Score 21; DB 13; Length 813;
 Best Local Similarity 58.3%; Pred. No. 6.6e+02;
 Matches 21; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 6 YGNTGGCCNTCCNTAYARGAYTAGGTNA 41
 :||| ||| ||| ||| :||| :||| :||| :|||
 Db 432 TGCATGGCCATCACCAACCATGATGAATTA 467

RESULT 82

CZ881566/c
 LOCUS 843 bp DNA linear GSS 27-JUL-2005
 DEFINITION OC_Ba0284F06.r OC_Ba Oryza coarctata genomic clone OC_Ba0284F06
 3', genomic survey sequence.

ACCESSION CZ881566
 VERSION
 KEYWORDS
 SOURCE GSS.

ORGANISM Oryza coarctata (Porteresia coarctata)

Oryza coarctata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
 Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
 Wing,R.

TITLE OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute

JOURNAL

COMMENT Unpublished (2005)
 Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 PCR Primers
 FORWARD: TAA TAC GAC TCA CTA TAG GG
 BACKWARD: CAC TCA TTA GGC ACC CCA
 Plate: 0284 row: F column: 06
 Seq primer: CAC TCA TTA GGC ACC CCA
 Class: BAC ends.

FEATURES

Location/Qualifiers

1..843
 /organism="Oryza coarctata"
 /mol_type="genomic DNA"
 /db_xref="taxon:77588"
 /clone="OC_Ba0284F06"
 /tissue_type="leaves"
 /dev_stage="mature"
 /lab_host="DH10B"
 /clone_lib="OC_Ba"
 /note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 63.3%; Score 21; DB 13; Length 843;
 Best Local Similarity 55.3%; Pred. No. 6.6e+02;
 Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 CAYGNTGGCCNTCCNTAYARGAYTAGGTNA 41
 ||| ||| ||| ||| :||| :||| :||| :|||
 Db 296 CATGCATGGTCATCACCCATCTAGATCATGAAGTCA 259

RESULT 83

BE548141
 LOCUS
 DEFINITION

ACCESSION BE548141

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 912)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM8449 row: p column: 24
 High quality sequence stop: 651.

FEATURES

source

1..912
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mrna"
 /db_xref="taxon:9606"
 /clone="IMAGE:3458567"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.4 Kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 63.3%; Score 21; DB 7; Length 912;
 Best Local Similarity 50.0%; Pred. No. 6.7e+02;
 Matches 21; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCTGGCCNTCCNTAYARGAYTAGGTNA 42

Db 617 GGGCCACCGCGGCGAGCCCTTACAAGACTATGAAGTCAG 658

RESULT 84

LOCUS

DEFINITION

BP701218 Osada Taira anterior neuroectoderm (ANE) PCS105 CDNA
 library Xenopus laevis cDNA clone XL500ml6x 5', mRNA sequence.

ACCESSION BP701218

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 369)
 Osada,S., Kitayama,A., Ueno,N. and Taira,M.
 Expression analysis of genes which are expressed in the anterior
 neuroectoderm of Xenopus embryos
 Unpublished (2004)
 Contact: Masanori Taira
 Department of Biological Sciences
 Graduate School of Science, University of Tokyo, Japan
 Science and Technology Corporation, Japan

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
 Tel: 81-03-5841-4434
 Fax: 81-03-5841-4434
 Email: m.taira@biol.s.u-tokyo.ac.jp,
 URL: <http://www.shigen.nig.ac.jp/nbrp/xenopus/est/>.
 Location/Qualifiers

FEATURES

source

1. .369
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL500ml6x"
 /tissue_type="anterior neuroectoderm"
 /dev_stage="late gastrula (stage 12.5)"
 /clone_lib="Osada Taira anterior neuroectoderm (ANE)
 pCS105 cDNA library"

ORIGIN

Query Match 62.7%; Score 20.8; DB 3; Length 369;
 Best Local Similarity 52.4%; Pred. No. 6.9e+02;
 Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
 |||||
 Db 219 GCTCATGCTTGCCTGCAACCCCTGGAACCAAGAAATGCGAAGTAAAG 260

RESULT 85

AV802339

LOCUS

DEFINITION AV802339 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-31-123 3',
 mRNA sequence.

ACCESSION AV802339

VERSION AV802339

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

REFERENCE 1 (bases 1 to 435)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)

JOURNAL

COMMENT

Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msek@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 site is in a modified pluscript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES

source

Location/Qualifiers
 1. .435
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL09-31-123"
 /dev_stage="plants at various developmental stages from
 germination to mature seeds"
 /lab_host="DH10B"
 /clone_lib="RAFL9"
 /notes="Site 1: BamHI; Site 2: SalI; subjected to
 dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
 hr) treatments"

ORIGIN

Query Match 62.7%; Score 20.8; DB 7; Length 435;
 Best Local Similarity 59.5%; Pred. No. 7.1e+02;
 Matches 22; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 4 CAYGCGTCCNTCNCNTAYAAAGAYTAYGARGTNA 40
 |||||
 Db 320 CATGCATTACCTCCCTCCCAACAAGAATTGGAAGTCA 356

RESULT 86

AZ177524

LOCUS

DEFINITION AZ177524 498 bp DNA linear GSS 30-AUG-2000
 SP 0147_Al H10 SP6E Strongylocentrotus purpuratus, purple sea
 urchin, sperm genomic BAC library Strongylocentrotus purpuratus
 genomic clone Plate=147 Col=19 Row=O, genomic survey sequence.

ACCESSION AZ177524

VERSION AZ177524

KEYWORDS GSS.

SOURCE Strongylocentrotus purpuratus

ORGANISM

Strongylocentrotus purpuratus
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidea; Euechinoidea; Echinacea; Echinoidea;
 Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 498)

AUTHORS

Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R.,
 Swartzell, S., Wallace, J.C., Pousetka, A.J., Livingston, B.T.,
 Wray, G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H.
 and Hood, L.

TITLE

A sea urchin genome project: Sequence scan, virtual map, and
 additional resources

JOURNAL

PUBMED

COMMENT

Contact: Cameron, RA, Davidson, EH, Hood, L
 Division of Biology 156-29
 California Institute of Technology
 Pasadena California 91125, USA
 Tel: (626) 395-8421
 Fax: (626) 793-3047
 Email: acameron@caltech.edu
 Plate: 147 row: O column: 19
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 498.

FEATURES

source

Location/Qualifiers
 1. .498
 /organism="Strongylocentrotus purpuratus"
 /mol_type="genomic DNA"
 /db_xref="taxon:7668"
 /clone="plate=147 Col=19 Row=O"
 /clone_lib="Strongylocentrotus purpuratus, purple sea
 urchin, sperm genomic BAC library"
 /notes="Organ: sperm; Vector: BACE3.6; BAC Clones in E-Coli
 DH10B"

ORIGIN

Query Match 62.7%; Score 20.8; DB 11; Length 498;
 Best Local Similarity 54.8%; Pred. No. 7.3e+02;
 Matches 23; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42

Db 53 GCAGACGCTGCCGCTCNACATTTAAGGACACCAATTATAG 94

RESULT 87

DN752781

LOCUS

DEFINITION DN752781 581 bp mRNA linear EST 01-APR-2005
 GLU-CF-9576 GLGC-LIB0001-cf Canis familiaris Normalized Mixed tissue
 cDNA Library Canis familiaris cDNA, mRNA sequence.

ACCESSION DN752781

VERSION DN752781.1


```

KEYWORDS
SOURCE  EST.
ORGANISM Canis familiaris (dog)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
          Canis.

REFERENCE
AUTHORS  1 (bases 1 to 581)
          Liu,Q.
TITLE    Direct Submission Gene Logic Inc
JOURNAL  Unpublished (2005)
COMMENT  Contact: Qing Liu
          Gene Logic Inc.
          610 Professional Drive, Gaithersburg, MD 20879, USA
          Tel: 301 987 1700
          Email: qlu@genelegic.com.

FEATURES
source
1..581
   /organism="Canis familiaris"
   /mol_type="mRNA"
   /strain="beagle"
   /db_xref="taxon:9615"
   /lab_host="EMDH10B"
   /clone_lib="GLGC-LTB0001-cf Canis familiaris Normalized
Mixed Tissue cDNA Library"
   /note="Organ: heart, liver, kidney, testis, and brain;
Vector: pCMVSPORT6.0; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match          62.7%; Score 20.8; DB 9; Length 581;
Best Local Similarity 55.0%; Pred. No. 7.5e+02;
Matches 22; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY  1 GCNCAYGCNTGCCNTCCNTAYAAAGAYTAYGARGTNA 40
    |||||
Db   339 GCCATGCGAGGCCCTCACACCACAGGCTACGCTCTTA 378

RESULT 88
DR3K23s/c
LOCUS    DR3K23S
DEFINITION Danio rerio genomic clone DKEY-3K23, genomic survey sequence.
ACCESSION AL740353
VERSION   AL740353.1 GI:213511134
KEYWORDS  GSS.
SOURCE    Danio rerio (zebrafish)
ORGANISM  Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS  1 (bases 1 to 590)
          Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE    Direct Submission
JOURNAL  Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
          Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
          humquerry@sanger.ac.uk Unpublished
COMMENT  This sequence was generated from the SP6 end of BAC 3K23. 3K23 is
          part of the Dariokey BAC Library created by R. Plasterk and N.V.
          Keygene.
          Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
source
1..590
   /organism="Danio rerio"
   /mol_type="genomic DNA"
   /db_xref="taxon:7955"
   /clone="DKEY-3K23"
   /tissue_type="testis"
   /note="Vector pIndigoBAC-536"

ORIGIN
Query Match          62.7%; Score 20.8; DB 14; Length 590;
Best Local Similarity 52.4%; Pred. No. 7.5e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY  1 GCNCAYGCNTGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db   95 GCCCATGTTGGATTTCCTCGTACAGAAAATGAATATAA 136

KEYWORDS
SOURCE  EST.
ORGANISM Canis familiaris (dog)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
          Canis.

REFERENCE
AUTHORS  1 (bases 1 to 647)
          Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
          Schmutz,J. and Myers,R.M.
TITLE    Expressed sequence tags from Gasterosteus aculeatus
JOURNAL  Unpublished (2003)
COMMENT  Contact: Kingsley, DM
          HHMI and Department of Developmental Biology
          Stanford University School of Medicine
          Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
          Tel: 650 725 5954
          Fax: 650 725 7739
          Email: kingsley@cmsgm.stanford.edu
          Plate: 23
          High quality sequence stop: 647.

FEATURES
source
1..647
   /organism="Gasterosteus aculeatus"
   /mol_type="mRNA"
   /strain="Salinas river, CA"
   /db_xref="taxon:69293"
   /clone="CDA23-H06"
   /sex="mixed male and female"
   /tissue_type="heads and internal organs combined"
   /dev_stage="adult"
   /clone_lib="SHGC-CDA"
   /note="Vector: lambda ZAP Express/pBK-CMV; Site_1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dT sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual pBK-CMV phagemid clones for
EST sequencing."

ORIGIN
Query Match          62.7%; Score 20.8; DB 5; Length 647;
Best Local Similarity 52.4%; Pred. No. 7.7e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY  1 GCNCAYGCNTGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db   95 GCCCATGTTGGATTTCCTCGTACAGAAAATGAATATAA 136

KEYWORDS
SOURCE  EST.
LOCUS    CW098794
DEFINITION 104_465_11003200_116_34373_056 Sorghum methylation filtered library
          CW098794
          669 bp DNA linear GSS 28-OCT-2004

ORIGIN
Query Match          62.7%; Score 20.8; DB 14; Length 590;
Best Local Similarity 52.4%; Pred. No. 7.5e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY  1 GCNCAYGCNTGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db   95 GCCCATGTTGGATTTCCTCGTACAGAAAATGAATATAA 136

```

(LibID: 104) Sorghum bicolor genomic clone 11003200, genomic survey sequence.

ACCESSION CW098794
VERSION CW098794.1 GI:54775524

KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)

ORGANISM Sorghum bicolor

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 669)

Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holean, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.

Sorghum genome sequencing by methylation filtration
PLoS Biol. 3 (1), e13 (2005)

15660154

COMMENT Contact: Bedell JA

Orion Genomics, LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 465 row: i column: 16

Seq primer: T3 Reverse

Class: methylation filtered

High quality sequence stop: 669.

Location/Qualifiers

FEATURES

source

1. .669

/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/cultivar="ATx623"

/db_xref="taxon:4558"

/clone_lib="11003200"

/clone_lib="Sorghum methylation filtered library (LibID: 104)"

/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 62.7%; Score 20.8; DB 13; Length 669;

Best Local Similarity 52.4%; Pred. No. 7.7e+02;

Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42

Db 416 GCCCGCGCTGGGCTTCGCGGTTCAACGATTCACAGAGAA 375

RESULT 91

CW309783

LOCUS

DEFINITION 104_798_11469373_148_36240_051 Sorghum methylation filtered library (LibID: 104) Sorghum bicolor genomic clone 11469373, genomic survey sequence.

ACCESSION CW309783

VERSION CW309783.1 GI:55025971

KEYWORDS GSS.

SOURCE Sorghum bicolor

ORGANISM Sorghum bicolor

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 677)

Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holean, H., Roe, B.A., Wiley, G., Korf, I.F.,

Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.

Sorghum genome sequencing by methylation filtration
PLoS Biol. 3 (1), e13 (2005)

15660154

COMMENT Contact: Bedell JA

Orion Genomics, LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 798 row: m column: 13

Seq primer: SWfor Forward

Class: methylation filtered

High quality sequence stop: 677.

Location/Qualifiers

FEATURES

source

1. .677

/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/cultivar="ATx623"

/db_xref="taxon:4558"

/clone_lib="11469373"

/clone_lib="Sorghum methylation filtered library (LibID: 104)"

/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 62.7%; Score 20.8; DB 13; Length 677;

Best Local Similarity 52.4%; Pred. No. 7.8e+02;

Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCNTGGCCNTCCNTAYARGAYTAYGARGTNAAR 42

Db 546 GCCCGCGCTGGGCTTCGCGGTTCAACGATTCACAGAGAA 587

RESULT 92

CW309782/c

LOCUS

DEFINITION 104_798_11469373_116_36239_051 Sorghum methylation filtered library (LibID: 104) Sorghum bicolor genomic clone 11469373, genomic survey sequence.

ACCESSION CW309782

VERSION CW309782.1 GI:55025970

KEYWORDS GSS.

SOURCE Sorghum bicolor

ORGANISM Sorghum bicolor

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 685)

Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holean, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.

Sorghum genome sequencing by methylation filtration
PLoS Biol. 3 (1), e13 (2005)

15660154

COMMENT Contact: Bedell JA

Orion Genomics, LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 798 row: m column: 13

Seq primer: T3 Reverse

Class: methylation filtered

```
High quality sequence stop: 685.
Location/Qualifiers
1. .685
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx823"
/db_xref="taxon:4558"
/clone="11469373"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."

ORIGIN
Query Match 62.7%; Score 20.8; DB 13; Length 685;
Best Local Similarity 52.4%; Pred. No. 7.8e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGCCTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 GCCCGCGTGGGTCTTCGCCGTTCAACGATTCACAAAGAAA 119

RESULT 93
LOCUS CA790172 861 bp mRNA linear EST 04-DEC-2002
DEFINITION AGENCOURT_10309392 NICHD_XGC_Emb1 Xenopus laevis cDNA clone
IMAGE:5156169 5', mRNA sequence.
ACCESSION CA790172
VERSION CA790172.1 GI:26035894
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 861)
/notes="Vector: pCMV-SF0RT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.55 kb. Constructed by Life Technologies. Note: This
is a Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 62.7%; Score 20.8; DB 4; Length 861;
Best Local Similarity 52.4%; Pred. No. 8.1e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGCCTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 GCCCGCGTGGGTCTTCGCCGTTCAACGATTCACAAAGAAA 119

High quality sequence stop: 685.
Location/Qualifiers
1. .685
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx823"
/db_xref="taxon:4558"
/clone="11469373"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."

ORIGIN
Query Match 62.7%; Score 20.8; DB 13; Length 685;
Best Local Similarity 52.4%; Pred. No. 7.8e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGCCTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 GCCCGCGTGGGTCTTCGCCGTTCAACGATTCACAAAGAAA 119

RESULT 94
LOCUS DW670971 906 bp mRNA linear EST 19-JAN-2006
DEFINITION CNB362-G10_Y1d-s SHGC-CNB2 Gasterosteus aculeatus cDNA clone
CNB362-G10 5', mRNA sequence.
ACCESSION DW670971
VERSION DW670971.1 GI:85469653
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 906)
/notes="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTTAGATCGAGCGCGCCG(T)25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria."

ORIGIN
Query Match 62.7%; Score 20.8; DB 10; Length 906;
Best Local Similarity 52.4%; Pred. No. 8.2e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGCCTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 GCCCGCGTGGGTCTTCGCCGTTCAACGATTCACAAAGAAA 119

FEATURES
source
1. .906
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CNB362-G10"
/tissue_type="adult"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CNB2"
/notes="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTTAGATCGAGCGCGCCG(T)25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria."

ORIGIN
Query Match 62.7%; Score 20.8; DB 10; Length 906;
Best Local Similarity 52.4%; Pred. No. 8.2e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGCCTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 GCCCGCGTGGGTCTTCGCCGTTCAACGATTCACAAAGAAA 119

FEATURES
source
1. .906
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CNB362-G10"
/tissue_type="adult"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CNB2"
/notes="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTTAGATCGAGCGCGCCG(T)25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria."

ORIGIN
Query Match 62.7%; Score 20.8; DB 10; Length 906;
Best Local Similarity 52.4%; Pred. No. 8.2e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGCCTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 GCCCGCGTGGGTCTTCGCCGTTCAACGATTCACAAAGAAA 119
```


LOCUS DN670301 1226 bp mRNA linear EST 29-MAR-2005
DEFINITION CFW55-G12.Y1d-s SHGC-CFW Gasterosteus aculeatus cDNA clone
CFW55-G12 5', mRNA sequence.

ACCESSION DN670301
VERSION DN670301.1 GI:61990358
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.

REFERENCE 1 (bases 1 to 1226)
AUTHORS Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schmutz,J. and Myers,R.M.

TITLE Expressed sequence tags from Gasterosteus aculeatus
JOURNAL Unpublished (2003)

COMMENT Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 55
High quality sequence start: 18
High quality sequence stop: 898.

FEATURES

source 1..1226
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CFW55-G12"
/sex="mixed male and female"
/tissue_type="gills"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CFW"
/notes="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTTCTAGATCGGACGCGCCCTT)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: <http://www.openbiosystems.com/cdna> library construction fa q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems: <http://www.openbiosystems.com/stickleback>"

ORIGIN

Query Match 62.7%; Score 20.8; DB 9; Length 1226;
Best Local Similarity 52.4%; Pred.No. 8.7e+02;
Matches 22; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
QY 1 GCNAYGCNTGCCNTCCNTAYAGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:
Db 727 GCCCATGTTTGATTTCTCGTACAGAAAAATGAATAAA 768

Search completed: June 10, 2006, 19:36:16
Job time : 5235.8 secs

THIS PAGE IS BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:21:38 ; Search time 94.15 Seconds
(without alignments)
834.695 Million cell updates/sec

Title: US-10-600-816-33

Perfect score: 33.2

Sequence: 1 gcnaygntgcctcnc.....ayaargaytaygargtnaar 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

Issued Patents NA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
- 2: /EMC Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
- 3: /EMC Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*
- 4: /EMC Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*
- 5: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
- 6: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
- 7: /EMC Celerra_SIDS3/ptodata/2/ina/PTUS COMB.seq:*
- 8: /EMC Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:*
- 9: /EMC Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
- 10: /EMC Celerra_SIDS3/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	90.4	601	3	US-09-949-016-29173
2	30	90.4	601	3	US-09-949-016-126957
3	30	90.4	1212	3	US-09-188-930-249
4	30	90.4	1212	3	US-09-312-283C-249
5	30	90.4	1228	3	US-09-495-050A-223
6	30	90.4	2289	3	US-09-949-016-3613
7	30	90.4	2456	3	US-09-949-016-572
8	30	90.4	26086	3	US-09-949-016-15355
9	30	90.4	26238	3	US-09-949-016-12314
10	20.4	61.4	1053	4	US-09-605-703B-1879
11	20	60.2	285986	3	US-09-949-016-12287
12	20	60.2	1230025	3	US-09-198-452A-1
13	20	60.2	1230230	3	US-09-438-185A-1
14	19.6	59.0	1005	3	US-09-248-796A-3066
15	19.4	58.4	2416	3	US-09-799-451-392
16	19.4	58.4	3738	3	US-09-543-681A-1420
17	19.2	57.8	1001	3	US-09-641-638-394
18	19.2	57.8	1001	3	US-10-170-097-394
19	19.2	57.8	1828	3	US-10-197-220-6
20	19.2	57.8	2496	3	US-09-976-594-637
21	19.2	57.8	10627	2	US-08-060-925A-12
22	19.2	57.8	12222	3	US-09-328-925-42
23	19.2	57.8	18141	3	US-09-949-016-14063

19	57.2	184	3	US-09-513-999C-21553	Sequence 21553, A
25	56.0	400	3	US-08-956-171E-2945	Sequence 2945, Ap
26	56.0	400	3	US-08-781-986A-2945	Sequence 2945, Ap
27	56.0	1305	3	US-09-540-236-933	Sequence 933, App
28	56.0	8190	3	US-09-949-016-17412	Sequence 17412, A
29	56.0	94750	3	US-09-596-002-38	Sequence 38, Appl
30	56.0	111677	3	US-09-949-016-16946	Sequence 16946, A
31	56.0	129908	3	US-09-585-858-1	Sequence 1, Appli
32	56.0	129908	3	US-10-270-878-1	Sequence 1, Appli
33	55.4	333	3	US-09-270-767-30495	Sequence 30495, A
34	55.4	1113	3	US-09-270-767-14344	Sequence 14344, A
35	55.4	6450	3	US-09-949-016-16596	Sequence 16596, A
36	55.4	13908	3	US-09-949-016-16736	Sequence 16736, A
37	55.4	19698	3	US-09-949-016-14079	Sequence 14079, A
38	55.4	23155	3	US-09-949-016-12366	Sequence 12366, A
39	55.4	35058	3	US-09-949-016-12607	Sequence 12607, A
40	55.4	35059	3	US-09-949-016-13831	Sequence 13831, A
41	55.4	45571	3	US-09-949-016-16262	Sequence 16262, A
42	55.4	81001	3	US-09-750-580-1	Sequence 1, Appli
43	55.4	209631	3	US-09-949-002-574	Sequence 574, App
44	55.4	209632	3	US-09-949-002-802	Sequence 802, App
45	55.4	260247	3	US-09-949-016-13358	Sequence 13358, A
46	54.8	171	3	US-08-991-789A-60	Sequence 60, Appl
47	54.8	171	3	US-09-062-451-60	Sequence 60, Appl
48	54.8	171	3	US-09-598-326-60	Sequence 60, Appl
49	54.8	171	3	US-09-289-198-60	Sequence 60, Appl
50	54.8	171	3	US-09-429-755-60	Sequence 60, Appl
51	54.8	171	3	US-09-699-295-60	Sequence 60, Appl
52	54.8	171	3	US-09-534-825A-60	Sequence 60, Appl
53	54.8	601	3	US-09-949-016-107890	Sequence 107890, A
54	54.8	669	3	US-09-302-540-1562	Sequence 1562, Ap
55	54.8	864	3	US-09-328-352-1269	Sequence 1269, Ap
56	54.8	1101	3	US-09-328-352-1336	Sequence 1336, Ap
57	54.8	1335	3	US-09-134-000C-2009	Sequence 2009, Ap
58	54.8	45418	3	US-09-949-016-15870	Sequence 15870, A
59	54.8	183770	3	US-09-949-016-15494	Sequence 15494, A
60	54.8	374159	3	US-09-949-016-15868	Sequence 15868, A
61	54.8	390890	3	US-09-949-016-14720	Sequence 14720, A
62	54.2	541	3	US-09-533-559-1061	Sequence 1061, Ap
63	54.2	601	3	US-09-949-016-151014	Sequence 151014, A
64	54.2	601	3	US-09-469-667-8	Sequence 8, Appli
65	54.2	878	2	US-08-469-667-8	Sequence 8, Appli
66	54.2	878	3	US-09-224-110-8	Sequence 8, Appli
67	54.2	878	3	US-09-988-292A-8	Sequence 8, Appli
68	54.2	878	7	PCT-US95-07289-8	Sequence 8, Appli
69	54.2	1512	3	US-09-016-434-850	Sequence 850, App
70	54.2	1659	4	US-10-044-205A-3	Sequence 3, Appli
71	54.2	1662	3	US-09-738-894A-1	Sequence 1, Appli
72	54.2	1662	3	US-09-802-117-1	Sequence 1, Appli
73	54.2	1662	3	US-09-964-469-1	Sequence 1, Appli
74	54.2	1662	3	US-10-217-745-1	Sequence 1, Appli
75	54.2	1662	5	US-10-425-962-1	Sequence 1, Appli
76	54.2	1953	3	US-09-583-110-560	Sequence 560, App
77	54.2	1953	3	US-09-107-433-912	Sequence 912, App
78	54.2	2148	3	US-09-248-796A-837	Sequence 837, App
79	54.2	2198	4	US-10-044-205A-1	Sequence 1, Appli
80	54.2	2249	3	US-09-802-117-5	Sequence 5, Appli
81	54.2	2249	3	US-10-217-745-5	Sequence 5, Appli
82	54.2	2371	2	US-08-757-541-4	Sequence 4, Appli
83	54.2	2371	3	US-09-033-275-4	Sequence 4, Appli
84	54.2	2371	3	US-09-342-581-4	Sequence 4, Appli
85	54.2	2745	3	US-09-623-624-5	Sequence 5, Appli
86	54.2	2745	3	US-10-270-595-5	Sequence 5, Appli
87	54.2	2763	3	US-09-685-403-1	Sequence 1, Appli
88	54.2	3007	3	US-09-193-562D-27	Sequence 27, Appl
89	54.2	3007	3	US-10-055-412B-27	Sequence 27, Appl
90	54.2	7172	3	US-08-961-527-120	Sequence 120, App
91	54.2	7411	3	US-10-001-189-67	Sequence 67, Appl
92	54.2	15566	3	US-09-949-016-17104	Sequence 17104, A
93	54.2	36851	3	US-09-738-894A-3	Sequence 3, Appli
94	54.2	36851	3	US-09-964-469-3	Sequence 3, Appli
95	54.2	36651	5	US-10-425-962-3	Sequence 3, Appli
96	54.2	100550	3	US-09-949-016-11835	Sequence 11835, A

C 97	18	54.2	100551	3	US-09-949-016-16207	Sequence 16207, A	C 170	17.6	53.0	181222	3	US-09-949-016-12956	Sequence 12956, A
C 98	18	54.2	137753	3	US-09-949-016-17404	Sequence 17404, A	C 171	17.6	53.0	21234	3	US-09-810-671-3	Sequence 3, Appli
C 99	18	54.2	144922	3	US-09-949-016-15980	Sequence 15980, A	C 172	17.6	53.0	21234	3	US-10-109-854-3	Sequence 3, Appli
C 100	18	54.2	239527	3	US-09-949-016-15980	Sequence 15980, A	C 173	17.6	53.0	21234	3	US-10-339-656-3	Sequence 3, Appli
C 101	18	54.2	268449	3	US-09-949-016-17244	Sequence 17244, A	C 174	17.6	53.0	23094	3	US-09-949-016-13468	Sequence 13468, A
C 102	17.8	53.6	294	3	US-09-313-294A-4456	Sequence 4456, Ap	C 175	17.6	53.0	24345	3	US-09-949-016-17549	Sequence 17549, A
C 103	17.8	53.6	359	3	US-09-270-767-3688	Sequence 3688, Ap	C 176	17.6	53.0	25190	3	US-09-949-016-15906	Sequence 15906, A
C 104	17.8	53.6	359	3	US-09-270-767-18970	Sequence 18970, A	C 177	17.6	53.0	36180	3	US-09-949-016-11745	Sequence 11745, A
C 105	17.8	53.6	645	3	US-09-989-534-1	Sequence 1, Appli	C 178	17.6	53.0	36181	3	US-09-949-016-16163	Sequence 16163, A
C 106	17.8	53.6	773	3	US-09-148-545-34	Sequence 34, Appl	C 179	17.6	53.0	36741	3	US-09-301-665-3	Sequence 3, Appli
C 107	17.8	53.6	773	3	US-09-621-011-34	Sequence 34, Appl	C 180	17.6	53.0	36741	3	US-09-782-378A-12	Sequence 12, Appl
C 108	17.8	53.6	1253	3	US-09-148-545-98	Sequence 98, Appl	C 181	17.6	53.0	40325	3	US-09-949-002-580	Sequence 580, App
C 109	17.8	53.6	1253	3	US-09-621-011-98	Sequence 98, Appl	C 182	17.6	53.0	40325	3	US-09-949-002-771	Sequence 771, App
C 110	17.8	53.6	1854	3	US-09-270-767-11900	Sequence 11900, A	C 183	17.6	53.0	41318	3	US-09-949-016-16225	Sequence 16225, A
C 111	17.8	53.6	1995	4	US-09-880-107-2276	Sequence 2276, Ap	C 184	17.6	53.0	49931	3	US-09-949-016-13727	Sequence 13727, A
C 112	17.8	53.6	2523	4	US-10-094-749-1277	Sequence 1277, Ap	C 185	17.6	53.0	49931	3	US-09-949-016-13728	Sequence 13728, A
C 113	17.8	53.6	2648	3	US-10-104-047-349	Sequence 349, App	C 186	17.6	53.0	49931	3	US-09-949-016-13729	Sequence 13729, A
C 114	17.8	53.6	4324	3	US-09-958-021-1	Sequence 1, Appli	C 187	17.6	53.0	54531	3	US-09-949-016-16267	Sequence 16267, A
C 115	17.8	53.6	5279	3	US-09-958-021-5	Sequence 5, Appli	C 188	17.6	53.0	54550	3	US-10-327-189-42	Sequence 42, Appl
C 116	17.8	53.6	5965	3	US-09-362-336A-1	Sequence 1, Appli	C 189	17.6	53.0	75212	3	US-09-949-016-13313	Sequence 13313, A
C 117	17.8	53.6	6028	3	US-09-362-336A-3	Sequence 3, Appli	C 190	17.6	53.0	75212	3	US-09-949-016-13314	Sequence 13314, A
C 118	17.8	53.6	8091	3	US-09-230-652-1	Sequence 1, Appli	C 191	17.6	53.0	75212	3	US-09-949-016-13315	Sequence 13315, A
C 119	17.8	53.6	8091	4	US-09-612-226B-1	Sequence 1, Appli	C 192	17.6	53.0	87523	3	US-09-949-016-12670	Sequence 12670, A
C 120	17.8	53.6	8357	3	US-09-484-970B-65	Sequence 65, Appl	C 193	17.6	53.0	87523	3	US-09-949-016-15047	Sequence 15047, A
C 121	17.8	53.6	16207	3	US-09-949-016-16747	Sequence 16747, A	C 194	17.6	53.0	87523	3	US-09-949-016-15048	Sequence 15048, A
C 122	17.8	53.6	27425	3	US-09-524-101D-18	Sequence 18, Appl	C 195	17.6	53.0	87523	3	US-09-949-016-15049	Sequence 15049, A
C 123	17.8	53.6	45842	3	US-09-949-016-16914	Sequence 16914, A	C 196	17.6	53.0	87869	3	US-09-949-016-11744	Sequence 11744, A
C 124	17.8	53.6	51770	3	US-09-949-016-13668	Sequence 13668, A	C 197	17.6	53.0	87869	3	US-09-949-016-15044	Sequence 15044, A
C 125	17.8	53.6	64137	3	US-09-949-016-14831	Sequence 14831, A	C 198	17.6	53.0	87869	3	US-09-949-016-15045	Sequence 15045, A
C 126	17.8	53.6	64171	3	US-09-949-016-12502	Sequence 12502, A	C 199	17.6	53.0	87869	3	US-09-949-016-15046	Sequence 15046, A
C 127	17.8	53.6	85222	3	US-09-949-016-14693	Sequence 14693, A	C 200	17.6	53.0	109925	3	US-09-949-016-13210	Sequence 13210, A
C 128	17.8	53.6	109250	3	US-09-949-016-12530	Sequence 12530, A	C 201	17.6	53.0	146480	3	US-09-949-016-15921	Sequence 15921, A
C 129	17.8	53.6	109251	3	US-09-949-016-17321	Sequence 17321, A	C 202	17.6	53.0	146039	3	US-09-949-016-12449	Sequence 12449, A
C 130	17.8	53.6	119214	3	US-09-949-016-12507	Sequence 12507, A	C 203	17.6	53.0	203475	3	US-09-949-016-14516	Sequence 14516, A
C 131	17.8	53.6	194537	3	US-09-949-016-12928	Sequence 12928, A	C 204	17.6	53.0	203475	3	US-09-949-016-14517	Sequence 14517, A
C 132	17.8	53.6	194933	3	US-09-949-016-14172	Sequence 14172, A	C 205	17.6	53.0	203475	3	US-09-949-016-14518	Sequence 14518, A
C 133	17.8	53.6	201529	3	US-09-949-016-14740	Sequence 14740, A	C 206	17.6	53.0	203475	3	US-09-949-016-14519	Sequence 14519, A
C 134	17.8	53.6	678533	3	US-09-949-016-14577	Sequence 14577, A	C 207	17.6	53.0	203475	3	US-09-949-016-17226	Sequence 17226, A
C 135	17.8	53.6	678533	3	US-09-949-016-14578	Sequence 14578, A	C 208	17.6	53.0	203475	3	US-09-949-016-17227	Sequence 17227, A
C 136	17.8	53.6	1830121	3	US-09-557-884-1	Sequence 1, Appli	C 209	17.6	53.0	203475	3	US-09-949-016-17228	Sequence 17228, A
C 137	17.8	53.6	1830121	3	US-09-643-990A-1	Sequence 1, Appli	C 210	17.6	53.0	203475	3	US-09-949-016-17229	Sequence 17229, A
C 138	17.8	53.6	1830121	3	US-10-158-865-1	Sequence 1, Appli	C 211	17.6	53.0	251769	3	US-09-949-016-13185	Sequence 13185, A
C 139	17.6	53.0	186	3	US-09-016-434-369	Sequence 369, App	C 212	17.6	53.0	251769	3	US-09-949-016-13186	Sequence 13186, A
C 140	17.6	53.0	260	3	US-09-313-294A-2362	Sequence 2362, Ap	C 213	17.6	53.0	266748	3	US-09-949-016-13187	Sequence 13187, A
C 141	17.6	53.0	273	3	US-09-313-294A-3713	Sequence 3713, Ap	C 214	17.6	53.0	266748	3	US-09-949-016-13188	Sequence 13188, A
C 142	17.6	53.0	296	3	US-09-313-294A-6899	Sequence 6899, Ap	C 215	17.6	53.0	267802	3	US-09-949-016-14543	Sequence 14543, A
C 143	17.6	53.0	393	5	US-09-543-679A-2843	Sequence 2843, Ap	C 216	17.6	53.0	421883	3	US-09-949-016-12557	Sequence 12557, A
C 144	17.6	53.0	601	3	US-09-949-016-95267	Sequence 95267, A	C 217	17.4	52.4	295	3	US-09-490-609B-13	Sequence 13, Appl
C 145	17.6	53.0	601	3	US-09-949-016-95268	Sequence 95268, A	C 218	17.4	52.4	461	3	US-09-621-976-16747	Sequence 16747, A
C 146	17.6	53.0	601	3	US-09-949-016-95445	Sequence 95445, A	C 219	17.4	52.4	601	3	US-09-949-016-31457	Sequence 31457, A
C 147	17.6	53.0	601	3	US-09-949-016-95446	Sequence 95446, A	C 220	17.4	52.4	601	3	US-09-949-016-31458	Sequence 31458, A
C 148	17.6	53.0	601	3	US-09-949-016-95623	Sequence 95623, A	C 221	17.4	52.4	601	3	US-09-949-016-31459	Sequence 31459, A
C 149	17.6	53.0	601	3	US-09-949-016-95624	Sequence 95624, A	C 222	17.4	52.4	601	3	US-09-949-016-120550	Sequence 120550, A
C 150	17.6	53.0	601	3	US-09-949-016-95801	Sequence 95801, A	C 223	17.4	52.4	601	3	US-09-949-016-127123	Sequence 127123, A
C 151	17.6	53.0	601	3	US-09-949-016-95802	Sequence 95802, A	C 224	17.4	52.4	601	3	US-09-949-016-127124	Sequence 127124, A
C 152	17.6	53.0	601	3	US-09-949-016-158400	Sequence 158400, A	C 225	17.4	52.4	601	3	US-09-949-016-127125	Sequence 127125, A
C 153	17.6	53.0	601	3	US-09-949-016-191560	Sequence 191560, A	C 226	17.4	52.4	601	3	US-09-949-016-127125	Sequence 127125, A
C 154	17.6	53.0	601	3	US-09-949-016-191561	Sequence 191561, A	C 227	17.4	52.4	618	3	US-09-893-737-91	Sequence 91, Appl
C 155	17.6	53.0	601	3	US-09-949-016-191738	Sequence 191738, A	C 228	17.4	52.4	652	3	US-09-669-751-8	Sequence 8, Appli
C 156	17.6	53.0	601	3	US-09-949-016-191739	Sequence 191739, A	C 229	17.4	52.4	828	5	US-09-974-300-1877	Sequence 1877, Ap
C 157	17.6	53.0	601	3	US-09-949-016-191916	Sequence 191916, A	C 230	17.4	52.4	927	3	US-09-489-039A-4427	Sequence 4427, Ap
C 158	17.6	53.0	601	3	US-09-949-016-191917	Sequence 191917, A	C 231	17.4	52.4	1143	3	US-09-480-297A-11	Sequence 11, Appl
C 159	17.6	53.0	601	3	US-09-949-016-192094	Sequence 192094, A	C 232	17.4	52.4	1143	5	US-10-366-791-11	Sequence 11, Appl
C 160	17.6	53.0	601	3	US-09-949-016-192095	Sequence 192095, A	C 233	17.4	52.4	1258	3	US-09-724-864-2	Sequence 2, Appli
C 161	17.6	53.0	601	3	US-09-949-016-205402	Sequence 205402, A	C 234	17.4	52.4	1284	3	US-09-134-001C-212	Sequence 212, App
C 162	17.6	53.0	700	3	US-09-735-271-737	Sequence 737, App	C 235	17.4	52.4	1574	3	US-09-614-221A-281	Sequence 281, App
C 163	17.6	53.0	700	3	US-09-735-271-738	Sequence 738, App	C 236	17.4	52.4	1710	4	US-10-094-749-1519	Sequence 1519, Ap
C 164	17.6	53.0	706	3	US-09-533-559-5577	Sequence 5577, Ap	C 237	17.4	52.4	3516	3	US-08-956-171B-235	Sequence 235, App
C 165	17.6	53.0	737	3	US-09-270-767-10478	Sequence 10478, A	C 238	17.4	52.4	3516	3	US-08-781-986A-235	Sequence 235, App
C 166	17.6	53.0	811	3	US-09-270-767-10591	Sequence 10591, A	C 239	17.4	52.4	6151	3	US-09-799-451-528	Sequence 528, App
C 167	17.6	53.0	3169	5	US-09-543-679A-2845	Sequence 2845, Ap	C 240	17.4	52.4	8280	5	US-09-936-271C-65	Sequence 65, Appl
C 168	17.6	53.0	4091	3	US-09-620-312D-339	Sequence 339, App	C 241	17.4	52.4	10080	5	US-09-936-271C-97	Sequence 97, Appl
C 169	17.6	53.0	15000	3	US-09-719-554-65	Sequence 65, Appl	C 242	17.4	52.4	17132	3	US-09-949-016-15361	Sequence 15361, A

C 243	17.4	52.4	19008	3	US-09-949-016-12923	Sequence 12923, A	C 316	17.2	51.8	11589	3	US-09-949-002-826	Sequence 826, App
C 244	17.4	52.4	24150	3	US-09-949-016-12438	Sequence 12438, A	C 317	17.2	51.8	13644	3	US-09-949-016-13054	Sequence 13054, A
C 245	17.4	52.4	83428	3	US-09-949-016-13610	Sequence 13610, A	C 318	17.2	51.8	28001	3	US-09-819-993-3	Sequence 3, Appli
C 246	17.4	52.4	130724	3	US-09-949-016-13753	Sequence 13753, A	C 319	17.2	51.8	28001	3	US-10-193-295-3	Sequence 3, Appli
C 247	17.4	52.4	133719	3	US-09-949-016-15092	Sequence 15092, A	C 320	17.2	51.8	28001	3	US-10-622-516-3	Sequence 3, Appli
C 248	17.4	52.4	138632	3	US-09-949-016-12781	Sequence 12781, A	C 321	17.2	51.8	36417	3	US-09-949-016-14755	Sequence 14755, A
C 249	17.4	52.4	198637	3	US-09-949-016-17393	Sequence 17393, A	C 322	17.2	51.8	46720	3	US-09-949-016-14129	Sequence 14129, A
C 250	17.4	52.4	235033	5	US-08-852-495C-1	Sequence 1, Appli	C 323	17.2	51.8	58133	3	US-09-949-016-16454	Sequence 16454, A
C 251	17.4	52.4	237326	5	US-08-852-495C-2	Sequence 2, Appli	C 324	17.2	51.8	77851	3	US-09-949-016-12508	Sequence 12508, A
C 252	17.4	52.4	246240	2	US-08-724-394A-20	Sequence 20, Appli	C 325	17.2	51.8	77867	3	US-09-949-016-13211	Sequence 13211, A
C 253	17.4	52.4	246240	2	US-08-724-394A-21	Sequence 21, Appli	C 326	17.2	51.8	77867	3	US-09-949-016-13212	Sequence 13212, A
C 254	17.4	52.4	246240	2	US-08-724-394A-22	Sequence 22, Appli	C 327	17.2	51.8	77940	3	US-09-949-016-12509	Sequence 12509, A
C 255	17.4	52.4	636591	3	US-09-949-016-11808	Sequence 11808, A	C 328	17.2	51.8	80004	3	US-09-949-016-16318	Sequence 16318, A
C 256	17.4	52.4	636591	3	US-09-949-016-13388	Sequence 13388, A	C 329	17.2	51.8	84761	3	US-09-949-016-11914	Sequence 11914, A
C 257	17.4	52.4	1830121	3	US-09-557-884-1	Sequence 1, Appli	C 330	17.2	51.8	84763	3	US-09-949-016-13914	Sequence 13914, A
C 258	17.4	52.4	1830121	3	US-09-643-990A-1	Sequence 1, Appli	C 331	17.2	51.8	84875	3	US-09-949-016-17334	Sequence 17334, A
C 259	17.4	52.4	1830121	3	US-10-158-865-1	Sequence 1, Appli	C 332	17.2	51.8	84875	3	US-09-949-016-17335	Sequence 17335, A
C 260	17.2	51.8	282	3	US-09-313-294A-1129	Sequence 1129, Ap	C 333	17.2	51.8	84875	3	US-09-949-016-17336	Sequence 17336, A
C 261	17.2	51.8	396	3	US-09-640-173-76	Sequence 76, Appli	C 334	17.2	51.8	84875	3	US-09-949-016-17337	Sequence 17337, A
C 262	17.2	51.8	396	3	US-09-713-550-76	Sequence 76, Appli	C 335	17.2	51.8	85122	3	US-09-949-016-14693	Sequence 14693, A
C 263	17.2	51.8	396	3	US-09-825-294-76	Sequence 76, Appli	C 336	17.2	51.8	85152	3	US-09-949-016-12665	Sequence 12665, A
C 264	17.2	51.8	396	3	US-09-970-966-76	Sequence 76, Appli	C 337	17.2	51.8	85152	3	US-09-949-016-12666	Sequence 12666, A
C 265	17.2	51.8	418	4	US-09-880-107-1417	Sequence 1417, Ap	C 338	17.2	51.8	85152	3	US-09-949-016-12667	Sequence 12667, A
C 266	17.2	51.8	514	3	US-08-956-171B-1033	Sequence 1033, Ap	C 339	17.2	51.8	85152	3	US-09-949-016-12668	Sequence 12668, A
C 267	17.2	51.8	514	3	US-08-781-986A-1033	Sequence 1033, Ap	C 340	17.2	51.8	88002	3	US-09-949-002-639	Sequence 639, App
C 268	17.2	51.8	601	3	US-09-949-016-21314	Sequence 21314, A	C 341	17.2	51.8	88002	3	US-09-949-002-717	Sequence 717, App
C 269	17.2	51.8	601	3	US-09-949-016-36095	Sequence 36095, A	C 342	17.2	51.8	101356	3	US-09-949-016-12364	Sequence 12364, A
C 270	17.2	51.8	601	3	US-09-949-016-36096	Sequence 36096, A	C 343	17.2	51.8	101357	3	US-09-949-016-16924	Sequence 16924, A
C 271	17.2	51.8	601	3	US-09-949-016-60272	Sequence 60272, A	C 344	17.2	51.8	106380	3	US-09-949-016-17553	Sequence 17553, A
C 272	17.2	51.8	601	3	US-09-949-016-60273	Sequence 60273, A	C 345	17.2	51.8	118382	3	US-09-949-016-15996	Sequence 15996, A
C 273	17.2	51.8	601	3	US-09-949-016-70975	Sequence 70975, A	C 346	17.2	51.8	118382	3	US-09-949-016-15997	Sequence 15997, A
C 274	17.2	51.8	601	3	US-09-949-016-74306	Sequence 74306, A	C 347	17.2	51.8	119214	3	US-09-949-016-12507	Sequence 12507, A
C 275	17.2	51.8	601	3	US-09-949-016-109622	Sequence 109622, A	C 348	17.2	51.8	128516	3	US-09-949-016-13501	Sequence 13501, A
C 276	17.2	51.8	601	3	US-09-949-016-109623	Sequence 109623, A	C 349	17.2	51.8	132456	3	US-09-949-016-13750	Sequence 13750, A
C 277	17.2	51.8	601	3	US-09-949-016-162919	Sequence 162919, A	C 350	17.2	51.8	132456	3	US-09-949-016-12541	Sequence 12541, A
C 278	17.2	51.8	601	3	US-09-949-016-197171	Sequence 197171, A	C 351	17.2	51.8	209513	3	US-09-949-016-17009	Sequence 17009, A
C 279	17.2	51.8	601	3	US-09-949-016-197172	Sequence 197172, A	C 352	17.2	51.8	209513	3	US-09-949-016-15094	Sequence 15094, A
C 280	17.2	51.8	601	3	US-09-949-016-197289	Sequence 197289, A	C 353	17.2	51.8	219964	3	US-09-949-016-15086	Sequence 15086, A
C 281	17.2	51.8	601	3	US-09-949-016-197290	Sequence 197290, A	C 354	17.2	51.8	276687	3	US-09-949-016-13840	Sequence 13840, A
C 282	17.2	51.8	601	3	US-09-949-016-197407	Sequence 197407, A	C 355	17.2	51.8	312957	3	US-09-949-001-31	Sequence 31, Appli
C 283	17.2	51.8	601	3	US-09-949-016-197408	Sequence 197408, A	C 356	17.2	51.8	312972	3	US-09-949-001-34	Sequence 34, Appli
C 284	17.2	51.8	601	3	US-09-949-016-197525	Sequence 197525, A	C 357	17.2	51.8	421118	3	US-09-949-016-16297	Sequence 16297, A
C 285	17.2	51.8	601	3	US-09-949-016-197526	Sequence 197526, A	C 358	17.2	51.8	1664976	3	US-08-916-421B-1	Sequence 1, Appli
C 286	17.2	51.8	601	3	US-09-949-002-2499	Sequence 2499, Ap	C 359	17.2	51.8	1664976	3	US-09-692-570-1	Sequence 1, Appli
C 287	17.2	51.8	601	3	US-09-879-792-34	Sequence 34, Appli	C 360	17.2	51.2	336	5	US-09-974-300-7425	Sequence 7425, Ap
C 288	17.2	51.8	751	3	US-08-868-288A-4	Sequence 4, Appli	C 361	17.2	51.2	412	3	US-08-513-999C-1486	Sequence 1486, Ap
C 289	17.2	51.8	1330	2	US-09-235-373-4	Sequence 4, Appli	C 362	17.2	51.2	528	3	US-09-615-192A-137	Sequence 137, App
C 290	17.2	51.8	1330	3	US-09-388-993-4	Sequence 4, Appli	C 363	17.2	51.2	528	3	US-09-169-789-137	Sequence 137, App
C 291	17.2	51.8	1330	3	US-09-501-714-4	Sequence 4, Appli	C 364	17.2	51.2	545	2	US-08-975-316-74	Sequence 74, Appli
C 292	17.2	51.8	1330	3	US-09-436-983-1	Sequence 1, Appli	C 365	17.2	51.2	545	3	US-09-615-192A-74	Sequence 74, Appli
C 293	17.2	51.8	2017	3	US-08-484-993B-9	Sequence 9, Appli	C 366	17.2	51.2	545	3	US-09-169-789-74	Sequence 74, Appli
C 294	17.2	51.8	2381	2	US-08-484-993B-9	Sequence 9, Appli	C 367	17.2	51.2	601	3	US-09-949-016-20869	Sequence 20869, A
C 295	17.2	51.8	2381	2	US-08-484-158B-9	Sequence 9, Appli	C 368	17.2	51.2	601	3	US-09-949-016-42273	Sequence 42273, A
C 296	17.2	51.8	2381	2	US-08-484-596A-9	Sequence 9, Appli	C 369	17.2	51.2	601	3	US-09-949-016-42340	Sequence 42340, A
C 297	17.2	51.8	2381	2	US-08-480-150A-9	Sequence 9, Appli	C 370	17.2	51.2	601	3	US-09-949-016-70530	Sequence 70530, A
C 298	17.2	51.8	2381	2	US-08-458-731-9	Sequence 9, Appli	C 371	17.2	51.2	601	3	US-09-949-016-73551	Sequence 73551, A
C 299	17.2	51.8	2381	3	US-08-149-223A-9	Sequence 9, Appli	C 372	17.2	51.2	601	3	US-09-949-016-73552	Sequence 73552, A
C 300	17.2	51.8	2463	3	US-09-583-110-1558	Sequence 1558, Ap	C 373	17.2	51.2	601	3	US-09-949-016-110216	Sequence 110216, A
C 301	17.2	51.8	2466	6	US-09-377-465A-1	Sequence 1, Appli	C 374	17.2	51.2	601	3	US-09-949-016-169374	Sequence 169374, A
C 302	17.2	51.8	2466	6	US-09-377-465A-3	Sequence 3, Appli	C 375	17.2	51.2	601	3	US-09-949-016-201190	Sequence 201190, A
C 303	17.2	51.8	2505	3	US-09-107-433-2157	Sequence 2157, Ap	C 376	17.2	51.2	720	3	US-09-710-279-3009	Sequence 3009, Ap
C 304	17.2	51.8	2869	3	US-10-104-047-1187	Sequence 1187, Ap	C 377	17.2	51.2	735	3	US-08-956-182-15	Sequence 15, Appli
C 305	17.2	51.8	3582	3	US-08-538-526-2	Sequence 2, Appli	C 378	17.2	51.2	738	3	US-09-893-525-35	Sequence 35, Appli
C 306	17.2	51.8	4074	3	US-09-016-434-1475	Sequence 1475, Ap	C 379	17.2	51.2	748	3	US-09-893-525-34	Sequence 34, Appli
C 307	17.2	51.8	4663	3	US-09-949-016-2387	Sequence 2387, Ap	C 380	17.2	51.2	775	3	US-08-961-083-17	Sequence 17, Appli
C 308	17.2	51.8	4965	3	US-08-961-527-143	Sequence 143, App	C 381	17.2	51.2	775	3	US-09-536-784-17	Sequence 17, Appli
C 309	17.2	51.8	6281	3	US-09-902-540-765	Sequence 765, App	C 382	17.2	51.2	775	3	US-09-765-271-17	Sequence 17, Appli
C 310	17.2	51.8	6908	3	US-09-620-3120-977	Sequence 977, App	C 383	17.2	51.2	775	3	US-09-765-272A-17	Sequence 17, Appli
C 311	17.2	51.8	6935	3	US-09-377-285B-21	Sequence 21, Appli	C 384	17.2	51.2	837	3	US-09-583-110-736	Sequence 736, App
C 312	17.2	51.8	6935	3	US-10-192-381-21	Sequence 21, Appli	C 385	17.2	51.2	861	3	US-09-107-433-1492	Sequence 1492, Ap
C 313	17.2	51.8	7953	3	US-09-949-016-16886	Sequence 16886, A	C 386	17.2	51.2	870	3	US-09-902-540-8534	Sequence 8534, Ap
C 314	17.2	51.8	9928	3	US-09-949-016-16885	Sequence 16885, A	C 387	17.2	51.2	873	3	US-09-543-681A-2289	Sequence 2289, Ap
C 315	17.2	51.8	11437	3	US-09-949-002-587	Sequence 587, App	C 388	17.2	51.2	1002	3	US-09-134-001C-1109	Sequence 1109, Ap

389	17	51.2	1024	3	US-09-328-475C-10	Sequence 10, Appl	462	16.8	50.6	601	3	US-09-949-016-78972	Sequence 78972, A
390	17	51.2	1026	3	US-09-348-796A-1037	Sequence 1037, Ap	463	16.8	50.6	601	3	US-09-949-016-78973	Sequence 78973, A
391	17	51.2	1395	3	US-09-252-991A-1954	Sequence 1954, Ap	464	16.8	50.6	601	3	US-09-949-016-78974	Sequence 78974, A
392	17	51.2	2001	3	US-09-270-767-31380	Sequence 31380, A	465	16.8	50.6	601	3	US-09-949-016-82164	Sequence 82164, A
393	17	51.2	2166	3	US-09-270-767-15080	Sequence 15080, A	466	16.8	50.6	601	3	US-09-949-016-127336	Sequence 127336, A
394	17	51.2	2692	3	US-10-104-047-1554	Sequence 1554, Ap	467	16.8	50.6	601	3	US-09-949-016-127337	Sequence 127337, A
395	17	51.2	3263	3	US-09-344-510B-6	Sequence 6, Appl	468	16.8	50.6	601	3	US-09-949-016-127338	Sequence 127338, A
396	17	51.2	3435	3	US-09-344-510B-7	Sequence 7, Appl	469	16.8	50.6	601	3	US-09-949-016-127339	Sequence 127339, A
397	17	51.2	3486	5	US-08-488-960-1	Sequence 1, Appl	470	16.8	50.6	601	3	US-09-949-016-132230	Sequence 132230, A
398	17	51.2	3572	3	US-09-710-279-3527	Sequence 3527, Ap	471	16.8	50.6	601	3	US-09-949-016-132231	Sequence 132231, A
399	17	51.2	3668	3	US-09-710-279-4046	Sequence 4046, Ap	472	16.8	50.6	601	3	US-09-949-016-152467	Sequence 152467, A
400	17	51.2	3740	3	US-09-949-016-2740	Sequence 2740, Ap	473	16.8	50.6	601	3	US-09-949-016-196716	Sequence 196716, A
401	17	51.2	3740	3	US-09-949-016-3726	Sequence 3726, Ap	474	16.8	50.6	601	3	US-09-949-016-205254	Sequence 205254, A
402	17	51.2	3743	3	US-09-949-016-514	Sequence 514, Ap	475	16.8	50.6	601	3	US-09-949-016-205255	Sequence 205255, A
403	17	51.2	3980	2	US-08-233-008A-1	Sequence 1, Appl	476	16.8	50.6	601	3	US-09-949-016-205256	Sequence 205256, A
404	17	51.2	3980	2	US-08-233-008A-5	Sequence 5, Appl	477	16.8	50.6	700	3	US-09-735-271-1288	Sequence 1288, Ap
405	17	51.2	4108	3	US-08-981-729-8	Sequence 8, Appl	478	16.8	50.6	700	3	US-09-735-271-1289	Sequence 1289, Ap
406	17	51.2	4108	3	US-08-981-446B-1	Sequence 1, Appl	479	16.8	50.6	707	3	US-09-641-638-302	Sequence 302, Ap
407	17	51.2	4108	3	US-08-613-811-8	Sequence 8, Appl	480	16.8	50.6	767	3	US-10-170-097-302	Sequence 302, Ap
408	17	51.2	5199	3	US-08-726-214-13	Sequence 13, Appl	481	16.8	50.6	807	3	US-09-248-796A-3092	Sequence 3092, Ap
409	17	51.2	5309	3	US-09-949-016-5172	Sequence 5172, Ap	482	16.8	50.6	844	3	US-09-533-559-7799	Sequence 7799, Ap
410	17	51.2	5390	3	US-09-993-525-41	Sequence 41, Appl	483	16.8	50.6	886	3	US-09-533-559-138	Sequence 138, Ap
411	17	51.2	5735	3	US-09-566-921-12	Sequence 12, Appl	484	16.8	50.6	1023	3	US-09-940-244-383	Sequence 383, Ap
412	17	51.2	6314	2	US-08-211-430-1	Sequence 1, Appl	485	16.8	50.6	1308	3	US-09-270-767-11389	Sequence 11389, A
413	17	51.2	6445	3	US-09-949-016-14019	Sequence 14019, A	486	16.8	50.6	1362	3	US-09-799-451-236	Sequence 236, Ap
414	17	51.2	6448	3	US-09-949-016-11922	Sequence 11922, A	487	16.8	50.6	1443	2	US-08-454-557C-13	Sequence 13, Appl
415	17	51.2	6519	2	US-08-233-008A-7	Sequence 7, Appl	488	16.8	50.6	1443	2	US-08-340-426D-13	Sequence 13, Appl
416	17	51.2	7740	3	US-09-949-016-14482	Sequence 14482, A	489	16.8	50.6	1443	2	US-08-450-673C-13	Sequence 13, Appl
417	17	51.2	8145	3	US-08-961-527-122	Sequence 122, Ap	490	16.8	50.6	1443	7	PCY-US95-17111A-13	Sequence 13, Appl
418	17	51.2	8345	3	US-09-949-016-16833	Sequence 16833, A	491	16.8	50.6	1797	3	US-09-482-273-69	Sequence 69, Appl
419	17	51.2	8518	3	US-09-302-540-899	Sequence 899, Ap	492	16.8	50.6	1802	3	US-09-949-016-5775	Sequence 5775, Ap
420	17	51.2	14082	3	US-09-949-016-11898	Sequence 11898, A	493	16.8	50.6	2059	3	US-10-104-047-1789	Sequence 1789, Ap
421	17	51.2	14087	3	US-09-949-016-14783	Sequence 14783, A	494	16.8	50.6	2119	4	US-10-094-749-579	Sequence 579, Ap
422	17	51.2	14326	3	US-09-949-016-14002	Sequence 14002, A	495	16.8	50.6	2359	2	US-08-426-169-5	Sequence 5, Appl
423	17	51.2	15402	3	US-09-949-016-12868	Sequence 12868, A	496	16.8	50.6	2359	3	US-09-233-813-5	Sequence 5, Appl
424	17	51.2	17842	3	US-09-949-016-13826	Sequence 13826, A	497	16.8	50.6	2359	7	PCY-US95-09470-5	Sequence 5, Appl
425	17	51.2	20116	3	US-09-949-016-16861	Sequence 16861, A	498	16.8	50.6	3259	3	US-10-104-047-438	Sequence 438, Ap
426	17	51.2	24333	3	US-09-639-207-9	Sequence 9, Appl	499	16.8	50.6	3668	2	US-08-849-345-1	Sequence 1, Appl
427	17	51.2	29291	3	US-09-949-016-16263	Sequence 16263, A	500	16.8	50.6	4059	3	US-09-799-451-560	Sequence 560, Ap
428	17	51.2	36228	3	US-09-949-016-12256	Sequence 12256, A	501	16.8	50.6	5058	3	US-09-889-595-1	Sequence 1, Appl
429	17	51.2	36228	3	US-09-949-016-15468	Sequence 15468, A	502	16.8	50.6	5058	3	US-09-899-595-1	Sequence 1, Appl
430	17	51.2	42610	3	US-09-949-016-13882	Sequence 13882, A	503	16.8	50.6	5631	3	US-09-949-016-15416	Sequence 15416, A
431	17	51.2	46492	3	US-09-949-016-12953	Sequence 12953, A	504	16.8	50.6	6718	3	US-09-949-016-4632	Sequence 4632, Ap
432	17	51.2	46492	3	US-09-949-016-12954	Sequence 12954, A	505	16.8	50.6	7571	3	US-09-949-016-12878	Sequence 12878, A
433	17	51.2	51336	3	US-09-949-016-12954	Sequence 16034, A	506	16.8	50.6	7571	3	US-09-949-016-17584	Sequence 17584, A
434	17	51.2	56516	2	US-08-596-306-1	Sequence 1, Appl	507	16.8	50.6	17037	3	US-09-949-016-16535	Sequence 16535, A
435	17	51.2	56516	3	US-09-338-907-1	Sequence 1, Appl	508	16.8	50.6	23501	3	US-09-949-016-17517	Sequence 17517, A
436	17	51.2	56516	3	US-09-218-207-1	Sequence 1, Appl	509	16.8	50.6	24154	3	US-09-949-016-16374	Sequence 16374, A
437	17	51.2	56520	3	US-09-338-907-179	Sequence 179, Ap	510	16.8	50.6	35913	3	US-09-949-016-16586	Sequence 16586, A
438	17	51.2	56520	3	US-09-218-207-179	Sequence 179, Ap	511	16.8	50.6	37030	3	US-08-311-731A-25	Sequence 25, Appl
439	17	51.2	65415	3	US-09-949-016-16699	Sequence 16699, A	512	16.8	50.6	40219	3	US-09-949-016-15337	Sequence 15337, A
440	17	51.2	105919	3	US-09-949-016-11769	Sequence 11769, A	513	16.8	50.6	41243	3	US-09-949-016-13006	Sequence 13006, A
441	17	51.2	129127	3	US-09-949-016-13481	Sequence 13481, A	514	16.8	50.6	41639	3	US-09-949-016-15471	Sequence 15471, A
442	17	51.2	177923	3	US-09-949-016-16513	Sequence 16513, A	515	16.8	50.6	46805	3	US-09-949-002-585	Sequence 585, Ap
443	17	51.2	219964	3	US-09-949-016-15086	Sequence 15086, A	516	16.8	50.6	46806	3	US-09-949-002-842	Sequence 842, Ap
444	17	51.2	235064	2	US-09-949-016-15390	Sequence 15390, A	517	16.8	50.6	53806	3	US-09-949-016-12572	Sequence 12572, A
445	16.8	50.6	58	2	US-08-784-512-1	Sequence 1, Appl	518	16.8	50.6	54550	3	US-10-327-189-42	Sequence 42, Appl
446	16.8	50.6	213	3	US-09-176-228-1	Sequence 1, Appl	519	16.8	50.6	59140	3	US-09-949-016-16023	Sequence 16023, A
447	16.8	50.6	213	3	US-09-248-796A-11898	Sequence 11898, A	520	16.8	50.6	60589	3	US-09-949-016-17070	Sequence 17070, A
448	16.8	50.6	237	3	US-09-270-767-24000	Sequence 24000, Ap	521	16.8	50.6	69324	3	US-09-949-016-15367	Sequence 15367, A
449	16.8	50.6	237	3	US-09-270-767-24000	Sequence 24000, A	522	16.8	50.6	86936	3	US-09-949-016-17314	Sequence 17314, A
450	16.8	50.6	278	3	US-09-270-767-26960	Sequence 26960, A	523	16.8	50.6	86945	3	US-09-949-016-13849	Sequence 13849, A
451	16.8	50.6	381	3	US-09-248-796A-8538	Sequence 8538, Ap	524	16.8	50.6	86945	3	US-09-949-016-13850	Sequence 13850, A
452	16.8	50.6	386	3	US-09-188-930-239	Sequence 239, Ap	525	16.8	50.6	86945	3	US-09-949-016-13851	Sequence 13851, A
453	16.8	50.6	386	3	US-09-312-283C-239	Sequence 239, Ap	526	16.8	50.6	86945	3	US-09-949-016-13852	Sequence 13852, A
454	16.8	50.6	514	2	US-08-967-820A-30	Sequence 30, Appl	527	16.8	50.6	86945	3	US-09-949-016-13853	Sequence 13853, A
455	16.8	50.6	601	3	US-09-949-016-18619	Sequence 18619, A	528	16.8	50.6	86945	3	US-09-949-016-13854	Sequence 13854, A
456	16.8	50.6	601	3	US-09-949-016-44566	Sequence 44566, A	529	16.8	50.6	86945	3	US-09-949-016-13855	Sequence 13855, A
457	16.8	50.6	601	3	US-09-949-016-48245	Sequence 48245, A	530	16.8	50.6	86945	3	US-09-949-016-13856	Sequence 13856, A
458	16.8	50.6	601	3	US-09-949-016-70820	Sequence 70820, A	531	16.8	50.6	86945	3	US-09-949-016-13857	Sequence 13857, A
459	16.8	50.6	601	3	US-09-949-016-73335	Sequence 73335, A	532	16.8	50.6	86945	3	US-09-949-016-13858	Sequence 13858, A
460	16.8	50.6	601	3	US-09-949-016-73336	Sequence 73336, A	533	16.8	50.6	88245	3	US-09-949-016-13935	Sequence 13935, A
461	16.8	50.6	601	3	US-09-949-016-78971	Sequence 78971, A	534	16.8	50.6	99370	3	US-09-949-016-12816	Sequence 12816, A

c 535	16.8	50.6	99370	3	US-09-949-016-17540	Sequence 17540, A	608	16.6	50.0	601	3	US-09-949-016-193176	Sequence 193176, A
c 536	16.8	50.6	102520	3	US-09-949-016-17367	Sequence 17367, A	609	16.6	50.0	601	3	US-09-949-016-193268	Sequence 193268, A
c 537	16.8	50.6	102526	3	US-09-949-016-12448	Sequence 12448, A	610	16.6	50.0	601	3	US-09-949-016-193269	Sequence 193269, A
c 538	16.8	50.6	147321	3	US-09-949-016-15450	Sequence 14550, A	611	16.6	50.0	601	3	US-09-949-016-193361	Sequence 193361, A
c 539	16.8	50.6	174029	3	US-09-949-016-12610	Sequence 12610, A	612	16.6	50.0	601	3	US-09-949-016-193362	Sequence 193362, A
c 540	16.8	50.6	174030	3	US-09-949-016-13880	Sequence 13880, A	613	16.6	50.0	601	3	US-09-949-016-205534	Sequence 205534, A
c 541	16.8	50.6	183112	3	US-09-949-016-14184	Sequence 14184, A	614	16.6	50.0	601	3	US-09-949-001-369	Sequence 369, App
c 542	16.8	50.6	190078	3	US-09-949-016-12707	Sequence 12707, A	615	16.6	50.0	601	3	US-09-949-001-370	Sequence 370, App
c 543	16.8	50.6	190078	3	US-09-949-016-17026	Sequence 17026, A	616	16.6	50.0	601	3	US-09-949-002-4713	Sequence 4713, App
c 544	16.8	50.6	221958	3	US-09-949-016-12173	Sequence 12173, A	617	16.6	50.0	601	3	US-09-949-002-4714	Sequence 4714, App
c 545	16.8	50.6	221966	3	US-09-949-016-15498	Sequence 15498, A	618	16.6	50.0	654	3	US-09-540-236-47	Sequence 47, Appl
c 546	16.8	50.6	250352	3	US-09-949-016-14724	Sequence 14724, A	619	16.6	50.0	758	3	US-09-902-540-1415	Sequence 1415, App
c 547	16.8	50.6	251769	3	US-09-949-016-13185	Sequence 13185, A	620	16.6	50.0	763	3	US-10-087-167-5	Sequence 5, Appl
c 548	16.8	50.6	251769	3	US-09-949-016-13186	Sequence 13186, A	621	16.6	50.0	902	3	US-09-602-777A-319	Sequence 319, App
c 549	16.8	50.6	266748	3	US-09-949-016-13187	Sequence 13187, A	622	16.6	50.0	927	4	US-09-605-703B-1281	Sequence 1281, App
c 550	16.8	50.6	266748	3	US-09-949-016-13188	Sequence 13188, A	623	16.6	50.0	927	4	US-09-605-703B-1283	Sequence 1283, App
c 551	16.8	50.6	346112	3	US-09-949-016-13185	Sequence 13185, A	624	16.6	50.0	1038	5	US-09-974-300-4698	Sequence 4698, App
c 552	16.8	50.6	784019	3	US-09-949-016-14033	Sequence 14033, A	625	16.6	50.0	1116	3	US-09-543-681A-3160	Sequence 3160, App
c 553	16.8	50.6	818128	3	US-09-949-016-14546	Sequence 14546, A	626	16.6	50.0	1248	3	US-09-248-796A-3364	Sequence 3364, App
c 554	16.8	50.6	818128	3	US-09-949-016-14547	Sequence 14547, A	627	16.6	50.0	1341	3	US-09-248-796A-3893	Sequence 3893, App
c 555	16.8	50.6	818128	3	US-09-949-016-14548	Sequence 14548, A	628	16.6	50.0	1407	3	US-09-543-681A-3892	Sequence 3892, App
c 556	16.8	50.6	818128	3	US-09-949-016-14549	Sequence 14549, A	629	16.6	50.0	1440	3	US-09-328-352-2335	Sequence 235, App
c 557	16.8	50.6	818128	3	US-09-949-016-14550	Sequence 14550, A	630	16.6	50.0	1500	3	US-10-087-167-67	Sequence 67, Appl
c 558	16.8	50.6	818128	3	US-09-949-016-14551	Sequence 14551, A	631	16.6	50.0	1515	3	US-10-087-167-75	Sequence 75, Appl
c 559	16.8	50.6	818128	3	US-09-949-016-14552	Sequence 14552, A	632	16.6	50.0	1737	3	US-09-710-092-15	Sequence 15, Appl
c 560	16.8	50.6	818128	3	US-09-949-016-14553	Sequence 14553, A	633	16.6	50.0	1737	3	US-10-231-354-15	Sequence 15, Appl
c 561	16.8	50.6	818128	3	US-09-949-016-14554	Sequence 14554, A	634	16.6	50.0	1767	3	US-10-087-167-120	Sequence 120, App
c 562	16.8	50.6	818128	3	US-09-949-016-14555	Sequence 14555, A	635	16.6	50.0	1782	3	US-10-087-167-124	Sequence 124, App
c 563	16.8	50.6	818128	3	US-09-949-016-14556	Sequence 14556, A	636	16.6	50.0	1820	2	US-08-757-046A-4	Sequence 4, Appl
c 564	16.8	50.6	818128	3	US-09-949-016-14557	Sequence 14557, A	637	16.6	50.0	1820	3	US-09-447-208-4	Sequence 4, Appl
c 565	16.8	50.6	818128	3	US-09-949-016-14558	Sequence 14558, A	638	16.6	50.0	1820	3	US-09-135-988-4	Sequence 4, Appl
c 566	16.8	50.6	818128	3	US-09-949-016-14559	Sequence 14559, A	639	16.6	50.0	1820	3	US-09-277-716-4	Sequence 4, Appl
c 567	16.8	50.6	818128	3	US-09-949-016-14560	Sequence 14560, A	640	16.6	50.0	1820	3	US-08-597-274A-4	Sequence 4, Appl
c 568	16.8	50.6	818128	3	US-09-949-016-14561	Sequence 14561, A	641	16.6	50.0	1820	3	US-08-908-909-4	Sequence 4, Appl
c 569	16.8	50.6	818128	3	US-09-949-016-14562	Sequence 14562, A	642	16.6	50.0	1820	3	US-09-609-161B-4	Sequence 4, Appl
c 570	16.8	50.6	818128	3	US-09-949-016-14563	Sequence 14563, A	643	16.6	50.0	1820	3	US-08-990-103-4	Sequence 4, Appl
c 571	16.8	50.6	818128	3	US-09-949-016-14564	Sequence 14564, A	644	16.6	50.0	1820	3	US-09-746-485A-4	Sequence 4, Appl
c 572	16.8	50.6	818128	3	US-09-949-016-14565	Sequence 14565, A	645	16.6	50.0	1820	3	US-10-126-139-4	Sequence 4, Appl
c 573	16.8	50.6	818128	3	US-09-949-016-14566	Sequence 14566, A	646	16.6	50.0	1820	3	US-10-126-798-4	Sequence 4, Appl
c 574	16.8	50.6	828152	3	US-09-949-016-12777	Sequence 12777, A	647	16.6	50.0	1820	3	US-10-126-777-4	Sequence 4, Appl
c 575	16.8	50.6	1664976	3	US-08-916-421B-1	Sequence 1, Appl	648	16.6	50.0	1823	2	US-08-680-726A-77	Sequence 77, Appl
c 576	16.8	50.6	1664976	3	US-09-692-570-1	Sequence 1, Appl	649	16.6	50.0	1823	2	US-08-680-726A-78	Sequence 78, Appl
c 577	16.8	50.6	4411529	3	US-09-103-840A-1	Sequence 1, Appl	650	16.6	50.0	1823	3	US-09-092-409-77	Sequence 77, Appl
c 578	16.6	50.0	178	3	US-09-513-999C-31548	Sequence 31548, A	651	16.6	50.0	1823	3	US-09-092-409-78	Sequence 78, Appl
c 579	16.6	50.0	184	2	US-08-485-657A-8	Sequence 8, Appl	652	16.6	50.0	1944	3	US-09-873-198-29	Sequence 29, Appl
c 580	16.6	50.0	184	3	US-09-366-380-8	Sequence 8, Appl	653	16.6	50.0	2052	4	US-10-094-749-796	Sequence 796, App
c 581	16.6	50.0	184	7	PCT-US95-02303-8	Sequence 8, Appl	654	16.6	50.0	2126	3	US-09-393-839-1	Sequence 1, Appl
c 582	16.6	50.0	262	3	US-09-313-294A-709	Sequence 709, App	655	16.6	50.0	2179	3	US-09-370-838-152	Sequence 152, App
c 583	16.6	50.0	276	3	US-09-513-999C-12351	Sequence 12351, A	656	16.6	50.0	2179	3	US-09-854-133-152	Sequence 152, App
c 584	16.6	50.0	313	3	US-09-513-999C-35864	Sequence 35864, A	657	16.6	50.0	2218	3	US-09-949-016-5231	Sequence 5231, App
c 585	16.6	50.0	314	4	US-09-880-107-1042	Sequence 1042, App	658	16.6	50.0	2356	3	US-09-221-017B-874	Sequence 874, App
c 586	16.6	50.0	486	3	US-09-543-681A-2354	Sequence 2354, App	659	16.6	50.0	2737	3	US-09-949-016-607	Sequence 607, App
c 587	16.6	50.0	502	3	US-09-621-876-18891	Sequence 18891, A	660	16.6	50.0	2854	3	US-09-799-451-42	Sequence 42, Appl
c 588	16.6	50.0	522	3	US-09-270-767-313	Sequence 313, App	661	16.6	50.0	3306	3	US-09-770-170-7	Sequence 7, Appl
c 589	16.6	50.0	522	3	US-09-270-767-15595	Sequence 15595, A	662	16.6	50.0	3306	3	US-09-614-221A-175	Sequence 175, App
c 590	16.6	50.0	522	3	US-09-902-540-2152	Sequence 2152, App	663	16.6	50.0	3822	3	US-09-849-747A-3	Sequence 3, Appl
c 591	16.6	50.0	531	3	US-09-854-133-303	Sequence 303, App	664	16.6	50.0	4281	3	US-09-710-092-13	Sequence 13, Appl
c 592	16.6	50.0	601	3	US-09-949-016-29771	Sequence 29771, A	665	16.6	50.0	4281	3	US-10-231-354-13	Sequence 13, Appl
c 593	16.6	50.0	601	3	US-09-949-016-34828	Sequence 34828, A	666	16.6	50.0	4281	4	US-10-154-419-68	Sequence 68, Appl
c 594	16.6	50.0	601	3	US-09-949-016-45058	Sequence 45058, A	667	16.6	50.0	5046	3	US-09-949-016-16731	Sequence 16731, A
c 595	16.6	50.0	601	3	US-09-949-016-52182	Sequence 52182, A	668	16.6	50.0	5189	3	US-09-949-016-14400	Sequence 14400, A
c 596	16.6	50.0	601	3	US-09-949-016-70994	Sequence 70994, A	669	16.6	50.0	5190	3	US-09-949-016-15342	Sequence 15342, A
c 597	16.6	50.0	601	3	US-09-949-016-73409	Sequence 73409, A	670	16.6	50.0	5247	3	US-09-949-016-14844	Sequence 14844, A
c 598	16.6	50.0	601	3	US-09-949-016-77543	Sequence 77543, A	671	16.6	50.0	5247	3	US-09-710-092-17	Sequence 17, Appl
c 599	16.6	50.0	601	3	US-09-949-016-126188	Sequence 126188, A	672	16.6	50.0	5958	3	US-10-231-354-17	Sequence 17, Appl
c 600	16.6	50.0	601	3	US-09-949-016-126189	Sequence 126189, A	673	16.6	50.0	6073	4	US-10-154-419-66	Sequence 66, Appl
c 601	16.6	50.0	601	3	US-09-949-016-142996	Sequence 142996, A	674	16.6	50.0	7158	3	US-09-543-681A-2132	Sequence 2132, App
c 602	16.6	50.0	601	3	US-09-949-016-176938	Sequence 176938, A	675	16.6	50.0	14485	3	US-09-876-216-3	Sequence 3, Appl
c 603	16.6	50.0	601	3	US-09-949-016-182425	Sequence 182425, A	676	16.6	50.0	14485	3	US-10-359-076-3	Sequence 3, Appl
c 604	16.6	50.0	601	3	US-09-949-016-182425	Sequence 182425, A	677	16.6	50.0	15379	3	US-09-949-016-14191	Sequence 14191, A
c 605	16.6	50.0	601	3	US-09-949-016-193082	Sequence 193082, A	678	16.6	50.0	15926	3	US-09-949-016-16973	Sequence 16973, A
c 606	16.6	50.0	601	3	US-09-949-016-193083	Sequence 193083, A	679	16.6	50.0	20634	3	US-09-949-016-16627	Sequence 16627, A
c 607	16.6	50.0	601	3	US-09-949-016-193175	Sequence 193175, A	680	16.6	50.0	23677	3	US-09-902-540-1218	Sequence 1218, App

C 681	16.6	50.0	24971	3	US-09-949-016-14444	Sequence 14444, A	754	16.4	49.4	601	3	US-09-949-016-73240	Sequence 73240, A
C 682	16.6	50.0	24972	3	US-09-949-016-15059	Sequence 15059, A	755	16.4	49.4	601	3	US-09-949-016-79355	Sequence 79355, A
C 683	16.6	50.0	26557	3	US-09-949-016-16791	Sequence 16791, A	756	16.4	49.4	601	3	US-09-949-016-79356	Sequence 79356, A
C 684	16.6	50.0	28198	3	US-09-949-016-12349	Sequence 12349, A	C 757	16.4	49.4	601	3	US-09-949-016-83556	Sequence 83556, A
C 685	16.6	50.0	31166	3	US-09-949-016-13464	Sequence 13464, A	C 758	16.4	49.4	601	3	US-09-949-016-83557	Sequence 83557, A
C 686	16.6	50.0	42610	3	US-09-949-016-13882	Sequence 13882, A	C 759	16.4	49.4	601	3	US-09-949-016-83582	Sequence 83582, A
C 687	16.6	50.0	48994	3	US-09-949-016-14091	Sequence 14091, A	C 760	16.4	49.4	601	3	US-09-949-016-83583	Sequence 83583, A
C 688	16.6	50.0	49301	3	US-09-949-016-16296	Sequence 16296, A	C 761	16.4	49.4	601	3	US-09-949-016-83608	Sequence 83608, A
C 689	16.6	50.0	50869	3	US-09-949-016-17598	Sequence 17598, A	C 762	16.4	49.4	601	3	US-09-949-016-83609	Sequence 83609, A
C 690	16.6	50.0	56976	3	US-09-949-016-17486	Sequence 17486, A	C 763	16.4	49.4	601	3	US-09-949-016-83634	Sequence 83634, A
C 691	16.6	50.0	57859	3	US-09-949-016-12334	Sequence 12334, A	C 764	16.4	49.4	601	3	US-09-949-016-83635	Sequence 83635, A
C 692	16.6	50.0	57859	3	US-09-949-016-14657	Sequence 14657, A	C 765	16.4	49.4	601	3	US-09-949-016-83660	Sequence 83660, A
C 693	16.6	50.0	57859	3	US-09-949-016-14658	Sequence 14658, A	C 766	16.4	49.4	601	3	US-09-949-016-83661	Sequence 83661, A
C 694	16.6	50.0	62311	3	US-09-949-016-14582	Sequence 14582, A	C 767	16.4	49.4	601	3	US-09-949-016-83886	Sequence 83886, A
C 695	16.6	50.0	62327	3	US-09-949-016-16809	Sequence 16809, A	C 768	16.4	49.4	601	3	US-09-949-016-83887	Sequence 83887, A
C 696	16.6	50.0	71251	3	US-09-949-016-15332	Sequence 15332, A	C 769	16.4	49.4	601	3	US-09-949-016-85074	Sequence 85074, A
C 697	16.6	50.0	72992	3	US-09-949-016-17592	Sequence 17592, A	C 770	16.4	49.4	601	3	US-09-949-016-85225	Sequence 85225, A
C 698	16.6	50.0	78125	3	US-09-949-016-16006	Sequence 16006, A	C 771	16.4	49.4	601	3	US-09-949-016-108514	Sequence 108514, A
C 699	16.6	50.0	86213	3	US-09-949-016-17240	Sequence 17240, A	C 772	16.4	49.4	601	3	US-09-949-016-112183	Sequence 112183, A
C 700	16.6	50.0	86213	3	US-09-949-016-17241	Sequence 17241, A	C 773	16.4	49.4	601	3	US-09-949-016-112224	Sequence 112224, A
C 701	16.6	50.0	86213	3	US-09-949-016-17242	Sequence 17242, A	C 774	16.4	49.4	601	3	US-09-949-016-145748	Sequence 145748, A
C 702	16.6	50.0	86213	3	US-09-949-016-17243	Sequence 17243, A	C 775	16.4	49.4	601	3	US-09-949-016-146016	Sequence 146016, A
C 703	16.6	50.0	100848	3	US-09-596-002-39	Sequence 39, Appl	C 776	16.4	49.4	601	3	US-09-949-016-146284	Sequence 146284, A
C 704	16.6	50.0	101674	3	US-09-949-016-12033	Sequence 12033, A	C 777	16.4	49.4	601	3	US-09-949-016-148208	Sequence 148208, A
C 705	16.6	50.0	101364	3	US-09-949-016-14433	Sequence 14433, A	C 778	16.4	49.4	601	3	US-09-949-016-148752	Sequence 148752, A
C 706	16.6	50.0	105055	3	US-09-949-016-14001	Sequence 14001, A	C 779	16.4	49.4	601	3	US-09-949-016-160380	Sequence 160380, A
C 707	16.6	50.0	105168	3	US-09-949-016-13296	Sequence 13296, A	C 780	16.4	49.4	601	3	US-09-949-016-160381	Sequence 160381, A
C 708	16.6	50.0	105189	3	US-09-949-016-13029	Sequence 13029, A	C 781	16.4	49.4	601	3	US-09-949-016-162251	Sequence 162251, A
C 709	16.6	50.0	126254	3	US-09-949-016-15341	Sequence 15341, A	C 782	16.4	49.4	601	3	US-09-949-016-162252	Sequence 162252, A
C 710	16.6	50.0	152070	3	US-09-949-016-15340	Sequence 15402, A	C 783	16.4	49.4	601	3	US-09-949-016-166836	Sequence 166836, A
C 711	16.6	50.0	153866	3	US-09-949-016-16919	Sequence 16919, A	C 784	16.4	49.4	601	3	US-09-949-016-173650	Sequence 173650, A
C 712	16.6	50.0	158735	3	US-09-949-016-11389	Sequence 11389, A	C 785	16.4	49.4	601	3	US-09-949-016-185360	Sequence 185360, A
C 713	16.6	50.0	158735	3	US-09-949-016-17130	Sequence 17130, A	C 786	16.4	49.4	601	3	US-09-949-016-185362	Sequence 185362, A
C 714	16.6	50.0	165651	3	US-09-949-016-13032	Sequence 13032, A	C 787	16.4	49.4	601	3	US-09-949-016-186436	Sequence 186436, A
C 715	16.6	50.0	178883	3	US-09-949-016-12733	Sequence 12733, A	C 788	16.4	49.4	601	3	US-09-949-016-187729	Sequence 187729, A
C 716	16.6	50.0	178884	3	US-09-949-016-13039	Sequence 13039, A	C 789	16.4	49.4	601	3	US-09-949-016-187770	Sequence 187770, A
C 717	16.6	50.0	188504	3	US-09-949-002-711	Sequence 711, Appl	C 790	16.4	49.4	601	3	US-09-949-016-190700	Sequence 190700, A
C 718	16.6	50.0	200918	3	US-09-949-002-647	Sequence 647, Appl	C 791	16.4	49.4	601	3	US-09-949-016-193122	Sequence 193122, A
C 719	16.6	50.0	237510	3	US-09-949-016-14873	Sequence 14273, A	C 792	16.4	49.4	615	4	US-09-605-703B-1549	Sequence 1549, Ap
C 720	16.6	50.0	245286	3	US-09-949-016-15497	Sequence 15497, A	C 793	16.4	49.4	647	3	US-09-533-559-2259	Sequence 2259, Ap
C 721	16.6	50.0	264206	3	US-09-949-016-12731	Sequence 12731, A	C 794	16.4	49.4	735	3	US-09-107-532A-1875	Sequence 1875, Ap
C 722	16.6	50.0	264304	3	US-09-949-016-13249	Sequence 13249, A	C 795	16.4	49.4	781	4	US-09-297-648-4011	Sequence 4011, Ap
C 723	16.6	50.0	265038	3	US-09-949-016-15779	Sequence 15779, A	C 796	16.4	49.4	804	3	US-09-134-000C-2801	Sequence 2801, Ap
C 724	16.6	50.0	271134	3	US-09-949-016-12705	Sequence 12705, A	C 797	16.4	49.4	828	3	US-09-302-540-9661	Sequence 9661, Ap
C 725	16.6	50.0	276687	3	US-09-949-016-13840	Sequence 13840, A	C 798	16.4	49.4	831	3	US-09-902-540-2034	Sequence 2034, Ap
C 726	16.6	50.0	294836	3	US-09-949-016-15974	Sequence 15974, A	C 799	16.4	49.4	921	3	US-09-252-991A-12978	Sequence 12978, A
C 727	16.6	50.0	305491	3	US-09-949-016-17550	Sequence 17550, A	C 800	16.4	49.4	972	3	US-09-328-352-3922	Sequence 3922, Ap
C 728	16.6	50.0	312957	3	US-09-949-001-31	Sequence 31, Appl	C 801	16.4	49.4	975	3	US-09-252-991A-12499	Sequence 12499, A
C 729	16.6	50.0	312972	3	US-09-949-001-34	Sequence 34, Appl	C 802	16.4	49.4	976	3	US-09-221-017B-1055	Sequence 1055, Ap
C 730	16.6	50.0	321022	3	US-09-949-016-11852	Sequence 11852, A	C 803	16.4	49.4	1005	3	US-09-270-767-14271	Sequence 14271, A
C 731	16.6	50.0	321022	3	US-09-949-016-14166	Sequence 14166, A	C 804	16.4	49.4	1014	3	US-09-543-681A-687	Sequence 687, App
C 732	16.6	50.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl	C 805	16.4	49.4	1045	3	US-09-221-017B-953	Sequence 953, App
C 733	16.6	50.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl	C 806	16.4	49.4	1191	4	US-09-528-225-36	Sequence 36, Appl
C 734	16.4	49.4	228	3	US-09-023-655-706	Sequence 706, Appl	C 807	16.4	49.4	1210	3	US-08-811-481-30	Sequence 30, Appl
C 735	16.4	49.4	279	3	US-09-313-2948-618	Sequence 618, Appl	C 808	16.4	49.4	1233	3	US-09-876-527-30	Sequence 30, Appl
C 736	16.4	49.4	347	3	US-09-401-064-356	Sequence 356, Appl	C 809	16.4	49.4	1248	3	US-09-489-039A-2907	Sequence 2907, Ap
C 737	16.4	49.4	430	3	US-09-270-767-30408	Sequence 30408, A	C 810	16.4	49.4	1348	3	US-08-948-113D-33	Sequence 33, Appl
C 738	16.4	49.4	468	5	US-09-974-300-7301	Sequence 7301, Ap	C 811	16.4	49.4	1362	3	US-10-197-220-147	Sequence 147, App
C 739	16.4	49.4	481	3	US-09-621-976-859	Sequence 859, Appl	C 812	16.4	49.4	1311	3	US-10-197-220-1	Sequence 1, Appl
C 740	16.4	49.4	490	3	US-09-613-9990-244	Sequence 244, Appl	C 813	16.4	49.4	1344	4	US-09-528-225-37	Sequence 37, Appl
C 741	16.4	49.4	555	3	US-09-680-889-5	Sequence 5, Appl	C 814	16.4	49.4	1413	2	US-08-239-276-9	Sequence 9, Appl
C 742	16.4	49.4	601	3	US-09-949-016-18179	Sequence 18179, A	C 815	16.4	49.4	1413	2	US-08-468-579B-9	Sequence 9, Appl
C 743	16.4	49.4	601	3	US-09-949-016-18180	Sequence 18180, A	C 816	16.4	49.4	1413	3	US-08-468-577B-9	Sequence 9, Appl
C 744	16.4	49.4	601	3	US-09-949-016-18908	Sequence 18908, A	C 817	16.4	49.4	1413	3	US-08-468-581A-9	Sequence 9, Appl
C 745	16.4	49.4	601	3	US-09-949-016-18909	Sequence 18909, A	C 818	16.4	49.4	1506	3	US-09-328-352-2814	Sequence 2814, Ap
C 746	16.4	49.4	601	3	US-09-949-016-20159	Sequence 20159, A	C 819	16.4	49.4	1521	3	US-09-489-039A-5902	Sequence 5902, Ap
C 747	16.4	49.4	601	3	US-09-949-016-44334	Sequence 44334, A	C 820	16.4	49.4	1336	3	US-09-248-796A-6713	Sequence 6713, Ap
C 748	16.4	49.4	601	3	US-09-949-016-44864	Sequence 44864, A	C 821	16.4	49.4	1879	3	US-09-949-016-4180	Sequence 4180, Ap
C 749	16.4	49.4	601	3	US-09-949-016-44865	Sequence 44865, A	C 822	16.4	49.4	1920	3	US-09-252-991A-7122	Sequence 7122, Ap
C 750	16.4	49.4	601	3	US-09-949-016-48597	Sequence 48597, A	C 823	16.4	49.4	1947	3	US-09-775-398-45	Sequence 45, Appl
C 751	16.4	49.4	601	3	US-09-949-016-56884	Sequence 56884, A	C 824	16.4	49.4	2119	3	US-09-167-206-1	Sequence 1, Appl
C 752	16.4	49.4	601	3	US-09-949-016-56885	Sequence 56885, A	C 825	16.4	49.4	2120	3	US-09-949-016-1958	Sequence 1958, Ap
C 753	16.4	49.4	601	3	US-09-949-016-72669	Sequence 72669, A	C 826	16.4	49.4	2160	3	US-09-252-991A-7089	Sequence 7089, Ap

C 827	16.4	49.4	2167	2	US-09-067-351-5	Sequence 5, Appli	C 900	16.4	49.4	44120	3	US-09-949-016-14153	Sequence 14153, A
C 828	16.4	49.4	2167	3	US-09-360-490-5	Sequence 5, Appli	C 901	16.4	49.4	44120	3	US-09-949-016-14154	Sequence 14154, A
C 829	16.4	49.4	2167	3	US-09-127-219B-2	Sequence 2, Appli	C 902	16.4	49.4	44120	3	US-09-949-016-14155	Sequence 14155, A
C 830	16.4	49.4	2274	3	US-09-949-016-212	Sequence 212, App	C 903	16.4	49.4	44120	3	US-09-949-016-14156	Sequence 14156, A
C 831	16.4	49.4	2281	3	US-10-104-047-254	Sequence 254, App	C 904	16.4	49.4	44998	3	US-09-949-016-12824	Sequence 12824, A
C 832	16.4	49.4	2288	3	US-09-221-017B-208	Sequence 208, App	C 905	16.4	49.4	44999	3	US-09-949-016-14832	Sequence 14832, A
C 833	16.4	49.4	2325	2	US-08-982-232-6	Sequence 6, Appli	C 906	16.4	49.4	44999	3	US-09-949-016-14833	Sequence 14833, A
C 834	16.4	49.4	2329	2	US-08-982-232-13	Sequence 13, Appli	C 907	16.4	49.4	44999	3	US-09-949-016-17108	Sequence 17108, A
C 835	16.4	49.4	2441	3	US-09-949-016-229	Sequence 29, Appli	C 908	16.4	49.4	44999	3	US-09-949-016-17109	Sequence 17109, A
C 836	16.4	49.4	2454	3	US-09-949-016-4491	Sequence 4491, Ap	C 909	16.4	49.4	46499	3	US-09-949-016-14032	Sequence 14032, A
C 837	16.4	49.4	2508	3	US-09-902-540-4526	Sequence 4526, Ap	C 910	16.4	49.4	50950	3	US-09-949-016-16559	Sequence 16559, A
C 838	16.4	49.4	2776	4	US-10-094-749-752	Sequence 752, App	C 911	16.4	49.4	56302	3	US-09-949-016-11892	Sequence 11892, A
C 839	16.4	49.4	2790	3	US-09-370-838-79	Sequence 79, Appli	C 912	16.4	49.4	57139	3	US-09-949-016-11771	Sequence 11771, A
C 840	16.4	49.4	2790	3	US-09-854-133-79	Sequence 79, Appli	C 913	16.4	49.4	57150	3	US-09-949-016-16233	Sequence 16233, A
C 841	16.4	49.4	2798	3	US-10-104-047-267	Sequence 267, App	C 914	16.4	49.4	57879	3	US-09-949-016-15932	Sequence 15932, A
C 842	16.4	49.4	2856	3	US-09-252-991A-10195	Sequence 10195, A	C 915	16.4	49.4	57879	3	US-09-949-016-16939	Sequence 16939, A
C 843	16.4	49.4	2878	3	US-09-949-016-5318	Sequence 5318, Ap	C 916	16.4	49.4	57879	3	US-09-949-016-16939	Sequence 16939, A
C 844	16.4	49.4	3037	3	US-09-949-016-5389	Sequence 5389, Ap	C 917	16.4	49.4	75394	3	US-09-949-016-11754	Sequence 11754, A
C 845	16.4	49.4	3311	2	US-08-239-276-10	Sequence 10, Appli	C 918	16.4	49.4	75394	3	US-09-949-016-12608	Sequence 12608, A
C 846	16.4	49.4	3311	2	US-08-468-579B-10	Sequence 10, Appli	C 919	16.4	49.4	81384	3	US-09-949-016-12422	Sequence 12422, A
C 847	16.4	49.4	3311	3	US-08-468-579B-10	Sequence 10, Appli	C 920	16.4	49.4	81384	3	US-09-949-016-15891	Sequence 15891, A
C 848	16.4	49.4	3311	3	US-08-468-579B-10	Sequence 10, Appli	C 921	16.4	49.4	82240	3	US-09-949-016-16279	Sequence 16279, A
C 849	16.4	49.4	3345	3	US-09-252-991A-8394	Sequence 10, Appli	C 922	16.4	49.4	82240	3	US-09-949-016-12065	Sequence 12065, A
C 850	16.4	49.4	3613	3	US-08-514-233A-1	Sequence 1, Appli	C 923	16.4	49.4	82240	3	US-09-949-016-14090	Sequence 14090, A
C 851	16.4	49.4	3613	3	US-09-015-399-9	Sequence 9, Appli	C 924	16.4	49.4	108441	3	US-09-949-016-14169	Sequence 14169, A
C 852	16.4	49.4	3637	3	US-09-566-921-5	Sequence 5, Appli	C 925	16.4	49.4	109159	3	US-09-949-016-14170	Sequence 14170, A
C 853	16.4	49.4	3678	3	US-09-252-991A-8447	Sequence 5, Appli	C 926	16.4	49.4	109159	3	US-09-949-016-14170	Sequence 14170, A
C 854	16.4	49.4	3893	3	US-09-221-017B-604	Sequence 604, App	C 927	16.4	49.4	125672	3	US-09-949-016-12245	Sequence 12245, A
C 855	16.4	49.4	4336	3	US-09-949-016-3781	Sequence 3781, Ap	C 928	16.4	49.4	131631	3	US-09-949-016-16956	Sequence 16956, A
C 856	16.4	49.4	4804	3	US-09-066-047-6	Sequence 6, Appli	C 929	16.4	49.4	131631	3	US-09-949-016-11757	Sequence 11757, A
C 857	16.4	49.4	4928	3	US-09-323-872A-32	Sequence 32, Appli	C 930	16.4	49.4	133358	3	US-09-949-016-16964	Sequence 16964, A
C 858	16.4	49.4	4928	3	US-09-072-433-17	Sequence 17, Appli	C 931	16.4	49.4	133358	3	US-09-949-016-12651	Sequence 12651, A
C 859	16.4	49.4	6316	4	US-09-880-107-1537	Sequence 1537, Ap	C 932	16.4	49.4	134434	3	US-09-949-016-17362	Sequence 17362, A
C 860	16.4	49.4	7893	3	US-09-949-016-885	Sequence 885, App	C 933	16.4	49.4	137394	3	US-09-949-016-13872	Sequence 13872, A
C 861	16.4	49.4	8607	3	US-09-949-016-13657	Sequence 13657, A	C 934	16.4	49.4	137394	3	US-09-949-016-12178	Sequence 12178, A
C 862	16.4	49.4	8608	3	US-09-949-016-12184	Sequence 12184, A	C 935	16.4	49.4	142504	3	US-09-949-016-13693	Sequence 13693, A
C 863	16.4	49.4	10464	3	US-09-957-974-1	Sequence 1, Appli	C 936	16.4	49.4	142504	3	US-09-949-016-12474	Sequence 12474, A
C 864	16.4	49.4	10970	3	US-08-716-351A-5	Sequence 5, Appli	C 937	16.4	49.4	146095	3	US-09-949-016-12872	Sequence 12872, A
C 865	16.4	49.4	12309	3	US-09-949-016-17042	Sequence 17042, A	C 938	16.4	49.4	146104	3	US-09-949-016-13239	Sequence 13239, A
C 866	16.4	49.4	12309	3	US-09-949-016-17043	Sequence 17043, A	C 939	16.4	49.4	146104	3	US-09-949-016-15970	Sequence 15970, A
C 867	16.4	49.4	12821	3	US-09-949-016-13589	Sequence 13589, A	C 940	16.4	49.4	149543	3	US-09-949-016-16089	Sequence 16089, A
C 868	16.4	49.4	13283	3	US-09-949-016-11858	Sequence 11858, A	C 941	16.4	49.4	175251	3	US-09-949-016-15970	Sequence 15970, A
C 869	16.4	49.4	13284	3	US-09-949-016-15900	Sequence 15900, A	C 942	16.4	49.4	181251	3	US-09-949-016-13125	Sequence 13125, A
C 870	16.4	49.4	13875	2	US-08-734-344-1	Sequence 1, Appli	C 943	16.4	49.4	186959	3	US-09-949-016-13125	Sequence 13125, A
C 871	16.4	49.4	15923	3	US-09-902-540-1095	Sequence 1095, Ap	C 944	16.4	49.4	186959	3	US-09-949-016-17321	Sequence 17321, A
C 872	16.4	49.4	16716	3	US-09-949-016-17060	Sequence 17060, A	C 945	16.4	49.4	191433	3	US-09-949-016-16144	Sequence 16144, A
C 873	16.4	49.4	16836	3	US-09-949-016-13700	Sequence 13700, A	C 946	16.4	49.4	202111	3	US-09-949-016-13877	Sequence 13877, A
C 874	16.4	49.4	17527	3	US-09-949-016-13145	Sequence 13145, A	C 947	16.4	49.4	205044	3	US-09-949-016-15851	Sequence 15851, A
C 875	16.4	49.4	18094	3	US-09-949-016-13179	Sequence 13179, A	C 948	16.4	49.4	205044	3	US-09-949-016-15853	Sequence 15853, A
C 876	16.4	49.4	20740	3	US-09-902-540-1223	Sequence 1223, Ap	C 949	16.4	49.4	205044	3	US-09-949-016-15419	Sequence 15419, A
C 877	16.4	49.4	24207	3	US-09-949-016-14364	Sequence 14364, A	C 950	16.4	49.4	212449	3	US-09-949-016-15419	Sequence 15419, A
C 878	16.4	49.4	24715	3	US-09-949-016-12979	Sequence 12979, A	C 951	16.4	49.4	223471	3	US-09-949-016-12387	Sequence 12387, A
C 879	16.4	49.4	27707	3	US-09-902-540-1226	Sequence 1226, Ap	C 952	16.4	49.4	223471	3	US-09-949-016-12724	Sequence 12724, A
C 880	16.4	49.4	28770	3	US-09-817-198C-3	Sequence 3, Appli	C 953	16.4	49.4	233129	3	US-09-949-016-16110	Sequence 16110, A
C 881	16.4	49.4	32616	3	US-09-949-016-16086	Sequence 16086, A	C 954	16.4	49.4	233129	3	US-08-852-495C-1	Sequence 1, Appli
C 882	16.4	49.4	32666	3	US-09-949-016-16086	Sequence 16086, A	C 955	16.4	49.4	235033	5	US-08-852-495C-2	Sequence 2, Appli
C 883	16.4	49.4	34011	3	US-09-949-016-12485	Sequence 12485, A	C 956	16.4	49.4	235033	5	US-08-852-495C-2	Sequence 2, Appli
C 884	16.4	49.4	37385	3	US-09-949-016-15354	Sequence 15354, A	C 957	16.4	49.4	246240	2	US-08-724-394A-20	Sequence 20, Appli
C 885	16.4	49.4	37996	3	US-09-949-016-16472	Sequence 16472, A	C 958	16.4	49.4	246240	2	US-08-724-394A-21	Sequence 21, Appli
C 886	16.4	49.4	38844	3	US-09-734-675-3	Sequence 3, Appli	C 959	16.4	49.4	246240	2	US-08-724-394A-22	Sequence 22, Appli
C 887	16.4	49.4	38844	3	US-10-060-333-3	Sequence 3, Appli	C 960	16.4	49.4	258775	3	US-09-949-016-16435	Sequence 16435, A
C 888	16.4	49.4	39498	3	US-09-949-016-12410	Sequence 12410, A	C 961	16.4	49.4	318184	3	US-09-949-016-12008	Sequence 12008, A
C 889	16.4	49.4	39498	3	US-09-949-016-16505	Sequence 16505, A	C 962	16.4	49.4	318184	3	US-09-949-016-17056	Sequence 17056, A
C 890	16.4	49.4	41199	3	US-09-949-016-17269	Sequence 17269, A	C 963	16.4	49.4	329200	3	US-10-027-983-11	Sequence 11, Appli
C 891	16.4	49.4	41522	3	US-09-949-016-17269	Sequence 17269, A	C 964	16.4	49.4	332000	3	US-09-949-016-14573	Sequence 14573, A
C 892	16.4	49.4	41523	3	US-09-949-016-15764	Sequence 15764, A	C 965	16.4	49.4	332000	3	US-09-949-016-14574	Sequence 14574, A
C 893	16.4	49.4	41589	3	US-09-949-016-12867	Sequence 12867, A	C 966	16.4	49.4	332000	3	US-09-531-120-209	Sequence 209, App
C 894	16.4	49.4	41589	3	US-09-949-016-12867	Sequence 12867, A	C 967	16.4	49.4	332000	3	US-09-949-016-11808	Sequence 11808, A
C 895	16.4	49.4	41863	3	US-09-949-016-15666	Sequence 15666, A	C 968	16.4	49.4	332000	3	US-09-949-016-13388	Sequence 13388, A
C 896	16.4	49.4	42693	3	US-09-949-016-17317	Sequence 17317, A	C 969	16.4	49.4	332000	3	US-09-949-016-12505	Sequence 12505, A
C 897	16.4	49.4	42693	3	US-09-949-016-17317	Sequence 17317, A	C 970	16.4	49.4	332000	3	US-09-949-016-12407	Sequence 12407, A
C 898	16.4	49.4	44120	3	US-09-949-016-14151	Sequence 14151, A	C 971	16.4	49.4	332000	3	US-09-949-016-12407	Sequence 12407, A
C 899	16.4	49.4	44120	3	US-09-949-016-14152	Sequence 14152, A	C 972	16.4	49.4	332000	3	US-09-949-016-17361	Sequence 17361, A
C 900	16.4	49.4	44120	3	US-09-949-016-14153	Sequence 14153, A	C 973	16.4	49.4	332000	3	US-09-949-016-14546	Sequence 14546, A


```
c 973 16.4 49.4 818128 3 US-09-949-016-14547
c 974 16.4 49.4 818128 3 US-09-949-016-14548
c 975 16.4 49.4 818128 3 US-09-949-016-14549
c 976 16.4 49.4 818128 3 US-09-949-016-14550
c 977 16.4 49.4 818128 3 US-09-949-016-14551
c 978 16.4 49.4 818128 3 US-09-949-016-14552
c 979 16.4 49.4 818128 3 US-09-949-016-14553
c 980 16.4 49.4 818128 3 US-09-949-016-14554
c 981 16.4 49.4 818128 3 US-09-949-016-14555
c 982 16.4 49.4 818128 3 US-09-949-016-14556
c 983 16.4 49.4 818128 3 US-09-949-016-14557
c 984 16.4 49.4 818128 3 US-09-949-016-14558
c 985 16.4 49.4 818128 3 US-09-949-016-14559
c 986 16.4 49.4 818128 3 US-09-949-016-14560
c 987 16.4 49.4 818128 3 US-09-949-016-14561
c 988 16.4 49.4 818128 3 US-09-949-016-14562
c 989 16.4 49.4 818128 3 US-09-949-016-14563
c 990 16.4 49.4 818128 3 US-09-949-016-14564
c 991 16.4 49.4 818128 3 US-09-949-016-14565
c 992 16.4 49.4 818128 3 US-09-949-016-14566
c 993 16.4 49.4 818128 3 US-09-949-016-14567
c 994 16.2 48.8 89 3 US-09-531-120-211
c 995 16.2 48.8 89 3 US-09-254-126D-22
c 996 16.2 48.8 234 3 US-09-540-236-1801
c 997 16.2 48.8 285 3 US-09-543-681A-2616
c 998 16.2 48.8 288 3 US-08-943-731-55
c 999 16.2 48.8 290 5 US-09-974-300-8070
1000 16.2 48.8 346 3 US-09-621-976-13443
1000 16.2 48.8 351 3 US-09-621-976-13420
```

ALIGNMENTS

```
RESULT 1
US-09-949-016-29173
; Sequence 29173, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29173
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29173
```

```
Query Match 90.4%; Score 30; DB 3; Length 601;
Best Local Similarity 64.3%; Pred. No. 0.00071;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 GCNCAYGNTGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 14 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 55
```

```
RESULT 2
US-09-949-016-126957
; Sequence 126957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126957
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126957
```

```
Query Match 90.4%; Score 30; DB 3; Length 601;
Best Local Similarity 64.3%; Pred. No. 0.00071;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 GCNCAYGNTGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 14 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 55
```

```
RESULT 3
US-09-188-930-249
; Sequence 249, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-188-930-249
```

```
Query Match 90.4%; Score 30; DB 3; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.00089;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 GCNCAYGNTGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 990 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1031
```

```
RESULT 4
US-09-312-283C-249
; Sequence 249, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
```

; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-249

Query Match 90.4%; Score 30; DB 3; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.00089;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCNTGCCNTCCNTAYAAAGAYTAYGARGCTNAAR 42
Db 990 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAG 1031

RESULT 5

US-09-495-050A-223
; Sequence 223, Application US/09495050A
; Patent No. 6492505

; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy

; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice

; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED
; FILE REFERENCE: PA-0013 US

; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999

; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program

; SEQ ID NO 223
; LENGTH: 1228
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6492505 2681738CT1
US-09-495-050A-223

Query Match 90.4%; Score 30; DB 3; Length 1228;
Best Local Similarity 64.3%; Pred. No. 0.00089;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCNTGCCNTCCNTAYAAAGAYTAYGARGCTNAAR 42
Db 541 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAG 582

RESULT 6

US-09-949-016-3613
; Sequence 3613, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3613
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3613

Query Match 90.4%; Score 30; DB 3; Length 2289;
Best Local Similarity 64.3%; Pred. No. 0.0011;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCNTGCCNTCCNTAYAAAGAYTAYGARGCTNAAR 42
Db 1117 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAG 1158

RESULT 7

US-09-949-016-572
; Sequence 572, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 572
; LENGTH: 2456
; TYPE: DNA

; ORGANISM: Human
US-09-949-016-572

Query Match 90.4%; Score 30; DB 3; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0011;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCNCNTGCCNTCCNTAYAAAGAYTAYGARGCTNAAR 42
Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAG 1312

RESULT 8

US-09-949-016-15355
; Sequence 15355, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15355
; LENGTH: 26086
; TYPE: DNA

; ORGANISM: Human
US-09-949-016-15355


```
; Query Match 58.4%; Score 19.4; DB 3; Length 2416;
; Best Local Similarity 55.6%; Pred. No. 85;
; Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GCNCGCCTGGCCNTCCNTAYARGATYAGAR 36
Db 1965 GCACACGATGGCTTGGCCAGACACAAACCAAGAG 1930

RESULT 16
US-09-543-681A-1420
; Sequence 1420, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1420
; LENGTH: 3738
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1420

Query Match 58.4%; Score 19.4; DB 3; Length 3738;
; Best Local Similarity 52.6%; Pred. No. 97;
; Matches 20; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 5 AYCCTGCCNTCCNTAYARGATYAGCTNAR 42
Db 3278 AAGCATGGTGCACACTTCAAGATTACAAAGTGAAA 3315

RESULT 17
US-09-641-638-394/c
; Sequence 394, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641.638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 394
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 479
; OTHER INFORMATION: 12-889-518 : polymorphic base G or A
; NAME/KEY: misc_binding
; LOCATION: 459..478
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 394
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 479
; OTHER INFORMATION: 12-889-518 : polymorphic base G or A
; NAME/KEY: misc_binding
; LOCATION: 459..478
; OTHER INFORMATION: 12-889-518.mis1, potential
```

```
; NAME/KEY: misc_binding
; LOCATION: 480..499
; OTHER INFORMATION: 12-889-518.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 976..995
; OTHER INFORMATION: upstream amplification primer, complement
; NAME/KEY: primer_bind
; LOCATION: 426..445
; OTHER INFORMATION: downstream amplification primer
; NAME/KEY: misc_binding
; LOCATION: 467..491
; OTHER INFORMATION: 12-889-518 potential probe
; US-09-641-638-394

Query Match 57.8%; Score 19.2; DB 3; Length 1001;
; Best Local Similarity 64.3%; Pred. No. 80;
; Matches 18; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 8 CNTGGCCNTCCNTAYARGATYAGA 35
Db 169 CATGCCATCTTCTTATAGGACAATGA 142

RESULT 18
US-10-170-097-394/c
; Sequence 394, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T14XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 394
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 479
; OTHER INFORMATION: 12-889-518 : polymorphic base G or A
; NAME/KEY: misc_binding
; LOCATION: 459..478
; OTHER INFORMATION: 12-889-518.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 480..499
; OTHER INFORMATION: 12-889-518.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 976..995
; OTHER INFORMATION: upstream amplification primer, complement
; NAME/KEY: primer_bind
; LOCATION: 426..445
; OTHER INFORMATION: downstream amplification primer
```

```
; ;
; NAME/KEY: misc_binding
; LOCATION: 467..491
; OTHER INFORMATION: 12-889-518 potential probe
US-10-170-097-394

Query Match
Best Local Similarity 57.8%; Score 19.2; DB 3; Length 1001;
Matches 18; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 8 CNTGGCCNTCCNCTAYAAARGAYTAYGA 35
Db 169 CATGGCCATCTTCTTATAGGACATGA 142

RESULT 19
US-10-197-220-6
; Sequence 6, Application US/10197220
; Patent No. 6919187
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 6
; LENGTH: 1828
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-197-220-6

Query Match
Best Local Similarity 57.8%; Score 19.2; DB 3; Length 1828;
Matches 21; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 6 YGNTGGCCNTCCNCTAYAAARGAYTAYGARGTNAAR 42
Db 1456 TGCGTGGCCGCTCTCCGTAGAAAGCACTTATGTATAA 1492

RESULT 20
US-09-976-594-637/c
; Sequence 637, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 637
; LENGTH: 2496
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 238877.8
US-09-976-594-637

Query Match
Best Local Similarity 57.8%; Score 19.2; DB 3; Length 2496;
Matches 21; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

; ;
; NAME/KEY: misc_binding
; LOCATION: 467..491
; OTHER INFORMATION: 12-889-518 potential probe
US-10-170-097-394

Query Match
Best Local Similarity 57.8%; Score 19.2; DB 3; Length 1001;
Matches 18; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 8 CNTGGCCNTCCNCTAYAAARGAYTAYGA 35
Db 169 CATGGCCATCTTCTTATAGGACATGA 142

RESULT 19
US-10-197-220-6
; Sequence 6, Application US/10197220
; Patent No. 6919187
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 6
; LENGTH: 1828
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-197-220-6

Query Match
Best Local Similarity 57.8%; Score 19.2; DB 3; Length 1828;
Matches 21; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 6 YGNTGGCCNTCCNCTAYAAARGAYTAYGARGTNAAR 42
Db 1456 TGCGTGGCCGCTCTCCGTAGAAAGCACTTATGTATAA 1492

RESULT 20
US-09-976-594-637/c
; Sequence 637, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 637
; LENGTH: 2496
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 238877.8
US-09-976-594-637

Query Match
Best Local Similarity 57.8%; Score 19.2; DB 2; Length 10627;
Matches 21; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCCNCTAYAAARGAYTAYGARGTNAAR 42
Db 830 GGGCATGCATCTTCTCCCTTCTATAGATGAGGAGGTTAAG 871

RESULT 22
US-09-328-925-42
; Sequence 42, Application US/09328925
; Patent No. 6610906
; GENERAL INFORMATION:
; APPLICANT: Kurachi, Kotoku
; APPLICANT: Kurachi, Sumiko
; TITLE OF INVENTION: Nucleotide Sequences for Gene Regulation and Methods of
; FILE REFERENCE: UM-03603
; CURRENT APPLICATION NUMBER: US/09/328,925
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: Patentin Ver. 2.0
```



```
; SEQ ID NO 42
; LENGTH: 12222
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-328-925-42

Query Match      57.8%; Score 19.2; DB 3; Length 12222;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 21; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYCNCNTGCCNCCNTAYAARGAYTAYGARGTNAAR 42
Db 2432 GGCATGCATCTTCTCCCTTCTATAGATGAGGAGGTTAAG 2473

RESULT 23
US-09-949-016-14063/c
; Sequence 14063, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14063
; LENGTH: 18141
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(18141)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14063

Query Match      57.8%; Score 19.2; DB 3; Length 18141;
Best Local Similarity 56.8%; Pred. No. 2e+02;
Matches 21; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 5 AYGCTGGCCNCCNCCNTAYAARGAYTAYGARGTNA 41
Db 12619 ACGCTGGACTTCTCAAAATTATGAGCATGAGTCAA 12583

RESULT 24
US-09-513-999C-21553/c
; Sequence 21553, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCES: 59 US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 21553
; LENGTH: 184
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-513-999C-21553

Query Match      57.2%; Score 19; DB 3; Length 184;
Best Local Similarity 63.3%; Pred. No. 58;
Matches 19; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYCNCNTGCCNCCNCCNTAYAARGAY 30
Db 44 GCGCAGCCTGGCCTTCCCAACACAGAT 15

RESULT 25
US-08-956-171E-2945
; Sequence 2945, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Gil H. Choi
; APPLICANT: Patrick S. Dillon
; APPLICANT: Craig A. Rosen
; APPLICANT: Steven C. Barash
; APPLICANT: Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2945:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2945:
US-08-956-171E-2945

Query Match      56.0%; Score 18.6; DB 3; Length 400;
Best Local Similarity 51.2%; Pred. No. 1.1e+02;
Matches 21; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYCNCNTGCCNCCNCCNTAYAARGAYTAYGARGTNA 41
Db 354 GCAGTGGCTGGCCAACTCGTTTTCAGGATTATCATCCAAA 394

RESULT 26
US-08-781-986A-2945
```

; Sequence 2945, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 2945:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 400 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-781-986A-2945

 Query Match 56.0%; Score 18.6; DB 3; Length 400;
 Best Local Similarity 51.2%; Pred. No. 1.1e+02;
 Matches 21; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

 QY 1 GCNCAYGCGTGGCCNTCCNTAYAAARGAYTAYGARGTNA 41
 DB 354 GCACGTGCTGCCCAACTCGTTTCAGGATTATCATCMA 394

 RESULT 27
 US-09-540-236-933
 ; Sequence 933, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540,236
 ; CURRENT FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 933
 ; LENGTH: 1305
 ; TYPE: DNA
 ; ORGANISM: M.catarrhalis
 ; US-09-540-236-933

 Query Match 56.0%; Score 18.6; DB 3; Length 1305;
 Best Local Similarity 51.2%; Pred. No. 1.6e+02;
 Matches 21; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

 QY 1 GCNCAYGCGTGGCCNTCCNTAYAAARGAYTAYGARGTNA 41

```
US-09-949-016-16946/c
; Sequence 16946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16946
; LENGTH: 111677
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16946

Query Match          56.0%; Score 18.6; DB 3; Length 111677;
Best Local Similarity 56.2%; Pred. No. 6.3e+02;
Matches 18; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy      2  CNCAYGNTGCCNTCNCNTAYAAARGAYTAY 33
Db      47626  CCCATGCCGGCCCTGCCCAAGACTTAC 47595

RESULT 31
US-09-585-858-1
; Sequence 1, Application US/09585858
; Patent No. 6492161
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arthor Aevansson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/09/585,858
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-09-585-858-1

Query Match          56.0%; Score 18.6; DB 3; Length 129908;
Best Local Similarity 56.1%; Pred. No. 6.6e+02;
Matches 18; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy      10  TGGCCNTCNCNTAYAAARGAYTAYGARGTNA 40
Db      28468  TGGCTCTCACCTTCAAGAGATGATGAAGTGA 28498

RESULT 32
US-10-270-878-1
; Sequence 1, Application US/10270878
; Patent No. 6818425
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
```

```
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arthor Aevansson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-878-1

Query Match          56.0%; Score 18.6; DB 3; Length 129908;
Best Local Similarity 58.1%; Pred. No. 6.6e+02;
Matches 18; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy      10  TGGCCNTCNCNTAYAAARGAYTAYGARGTNA 40
Db      28468  TGGCTCTCACCTTCAAGAGATGATGAAGTGA 28498

RESULT 33
US-09-270-767-30495/c
; Sequence 30495, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30495
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30495

Query Match          55.4%; Score 18.4; DB 3; Length 333;
Best Local Similarity 54.3%; Pred. No. 1.3e+02;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy      2  CNCAYGNTGCCNTCNCNTAYAAARGAYTAYGAR 36
Db      226  CGCATTGCTTGTGTCGCCCGTATAAGACTAGGAA 192

RESULT 34
US-09-270-767-14344/c
; Sequence 14344, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14344
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14344
```

```
Query Match          55.4%; Score 18.4; DB 3; Length 1113;
Best Local Similarity 54.3%; Pred. No. 1.9e+02;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY      2  CMCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGAR 36
Db      1006  CGCATTCGTTGTGTGCGCGTATAAAGACTAGGAA 972

RESULT 35
US-09-949-016-16596
; Sequence 16596, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16596
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16596

Query Match          55.4%; Score 18.4; DB 3; Length 6450;
Best Local Similarity 52.4%; Pred. No. 3.3e+02;
Matches 22; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY      1  GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db      1728  GCAAAAGCTTGGCAGGCGCTTCAAGGATTTGGAGGTACAG 1769

RESULT 36
US-09-949-016-16726
; Sequence 16726, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16726
; LENGTH: 13908
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16726

Query Match          55.4%; Score 18.4; DB 3; Length 13908;
Best Local Similarity 54.3%; Pred. No. 4.2e+02;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY      8  CNTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db      1728  GCAAAAGCTTGGCAGGCGCTTCAAGGATTTGGAGGTACAG 1769

Query Match          55.4%; Score 18.4; DB 3; Length 19698;
Best Local Similarity 51.4%; Pred. No. 4.7e+02;
Matches 19; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY      6  YGCTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db      18964  TGGTTCACCTTCCCTTTTAAAGATCTCTATGTATAA 18928

RESULT 37
US-09-949-016-14079/c
; Sequence 14079, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14079
; LENGTH: 19698
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14079

Query Match          55.4%; Score 18.4; DB 3; Length 19698;
Best Local Similarity 51.4%; Pred. No. 4.7e+02;
Matches 19; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY      6  YGCTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db      18964  TGGTTCACCTTCCCTTTTAAAGATCTCTATGTATAA 18928

RESULT 38
US-09-949-016-12366
; Sequence 12366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12366
; LENGTH: 23155
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23155)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12366

Query Match          55.4%; Score 18.4; DB 3; Length 23155;
Best Local Similarity 52.4%; Pred. No. 4.9e+02;
Matches 22; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY      1  GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db      1728  GCAAAAGCTTGGCAGGCGCTTCAAGGATTTGGAGGTACAG 1769
```

Db 17278 GCAAAAGCTTGGCCAGGCGCTTCAAGGATTGGAGGTACAG 17319

RESULT 39

US-09-949-016-12607
; Sequence 12607, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12607
; LENGTH: 35058
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12607

Query Match 55.4%; Score 18.4; DB 3; Length 35058;
Best Local Similarity 48.7%; Pred. No. 5.6e+02;
Matches 19; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 4 CAYCNGTGGCCNTCNCNTAYARGAYTAYGARGTNAAR 42

Db 23763 CATACATCCTCCTCAGCTTACAAAGATGACGGGATTAAAG 23801

RESULT 40

US-09-949-016-13831
; Sequence 13831, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13831
; LENGTH: 35059
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13831

Query Match 55.4%; Score 18.4; DB 3; Length 35059;
Best Local Similarity 48.7%; Pred. No. 5.6e+02;
Matches 19; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 4 CAYCNGTGGCCNTCNCNTAYARGAYTAYGARGTNAAR 42

Db 23763 CATACATCCTCCTCAGCTTACAAAGATGACGGGATTAAAG 23801

RESULT 41

US-09-949-016-16262

; Sequence 16262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16262
; LENGTH: 45571
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16262

Query Match 55.4%; Score 18.4; DB 3; Length 45571;
Best Local Similarity 59.4%; Pred. No. 6e+02;
Matches 19; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 4 CAYCNGTGGCCNTCNCNTAYARGAYTAYGA 35

Db 18161 CATGCTGGCCAAATTCATATAAATTTAGGA 18192

RESULT 42

US-09-750-580-1
; Sequence 1, Application US/09750580
; Patent No. 6455280
; GENERAL INFORMATION:
; APPLICANT: Yen, Frances
; APPLICANT: Denison, Blake
; APPLICANT: Bour, Barbara
; APPLICANT: Bihain, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Salter-Cid, Luisa
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
; FILE REFERENCE: 89.US2.CIP
; CURRENT APPLICATION NUMBER: US/09/750,580
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 09/599,362
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB00/0101
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB99/02058
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 49/469/099
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 60/141,032
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 81001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10946..12946
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon

```

; LOCATION: 12947..12958
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 13470..13526
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 13641..13752
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 14271..15968
; OTHER INFORMATION: exon 4
; NAME/KEY: misc feature
; LOCATION: 15969..17969
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1239
; OTHER INFORMATION: 20-828-311 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12347
; OTHER INFORMATION: 17-42-319 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 15241
; OTHER INFORMATION: 17-41-250 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 42218
; OTHER INFORMATION: 20-841-149 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 45442
; OTHER INFORMATION: 20-842-115 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 77058
; OTHER INFORMATION: 20-853-415 : polymorphic base C or T
; NAME/KEY: primer_bind
; LOCATION: 929..949
; OTHER INFORMATION: 20-828.pu
; NAME/KEY: primer_bind
; LOCATION: 1357..1377
; OTHER INFORMATION: 20-828.rp complement
; NAME/KEY: primer_bind
; LOCATION: 12029..12050
; OTHER INFORMATION: 17-42.pu
; NAME/KEY: primer_bind
; LOCATION: 12581..12603
; OTHER INFORMATION: 17-42.rp complement
; NAME/KEY: primer_bind
; LOCATION: 14992..15012
; OTHER INFORMATION: 17-41.pu
; NAME/KEY: primer_bind
; LOCATION: 15460..15482
; OTHER INFORMATION: 17-41.rp complement
; NAME/KEY: primer_bind
; LOCATION: 42070..42090
; OTHER INFORMATION: 20-841.pu
; NAME/KEY: primer_bind
; LOCATION: 42572..42591
; OTHER INFORMATION: 20-841.rp complement
; NAME/KEY: primer_bind
; LOCATION: 45328..45347
; OTHER INFORMATION: 20-842.pu
; NAME/KEY: primer_bind
; LOCATION: 45863..45883
; OTHER INFORMATION: 20-842.rp complement
; NAME/KEY: primer_bind
; LOCATION: 76644..76664
; OTHER INFORMATION: 20-853.pu
; NAME/KEY: primer_bind
; LOCATION: 77166..77185
; OTHER INFORMATION: 20-853.rp complement
; NAME/KEY: primer_bind
; LOCATION: 1220..1238
; OTHER INFORMATION: 20-828-311.mis
; NAME/KEY: primer_bind
; LOCATION: 1240..1258

```

```

; OTHER INFORMATION: 20-828-311.mis complement
; NAME/KEY: primer_bind
; LOCATION: 12328..12346
; OTHER INFORMATION: 17-42-319.mis
; NAME/KEY: primer_bind
; LOCATION: 12348..12366
; OTHER INFORMATION: 17-42-319.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15222..15240
; OTHER INFORMATION: 17-41-250.mis
; NAME/KEY: primer_bind
; LOCATION: 15242..15260
; OTHER INFORMATION: 17-41-250.mis complement
; NAME/KEY: primer_bind
; LOCATION: 42199..42217
; OTHER INFORMATION: 20-841-149.mis
; NAME/KEY: primer_bind
; LOCATION: 42219..42237
; OTHER INFORMATION: 20-841-149.mis complement
; NAME/KEY: primer_bind
; LOCATION: 45423..45441
; OTHER INFORMATION: 20-842-115.mis
; NAME/KEY: primer_bind
; LOCATION: 45443..45461
; OTHER INFORMATION: 20-842-115.mis complement
; NAME/KEY: primer_bind
; LOCATION: 77039..77057
; OTHER INFORMATION: 20-853-415.mis
; NAME/KEY: primer_bind
; LOCATION: 77059..77077
; OTHER INFORMATION: 20-853-415.mis complement
; NAME/KEY: misc binding
; LOCATION: 1227..1251
; OTHER INFORMATION: 20-828-311.probe
; NAME/KEY: misc binding
; LOCATION: 12335..12359
; OTHER INFORMATION: 17-42-319.probe
; NAME/KEY: misc binding
; LOCATION: 15229..15253
; OTHER INFORMATION: 17-41-250.probe
; NAME/KEY: misc binding
; LOCATION: 42206..42230
; OTHER INFORMATION: 20-841-149.probe
; NAME/KEY: misc binding
; LOCATION: 45430..45454
; OTHER INFORMATION: 20-842-115.probe
; NAME/KEY: misc binding
; LOCATION: 77046..77070
; OTHER INFORMATION: 20-853-415.probe
; US-09-750-580-1

```

```

Query Match      55.4%; Score 18.4; DB 3; Length 81001;
Best Local Similarity 54.3%; Pred. No. 7.1e+02;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```

```

QY      8  CMTGGCCNTCCNCTAYAAAGAYTAYGARGTNAAR 42
      |||||
Db      18627  CTTGGCTTCTCATTTTACAGATGAGGAAATTAAG 18661

```

```

RESULT 43
US-09-949-002-574/c
; Sequence 574, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401

```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 574
; LENGTH: 209631
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(209631)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-574

Query Match      55.4%; Score 18.4; DB 3; Length 209631;
Best Local Similarity 59.3%; Pred. No. 9.2e+02;
Matches 16; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy      16 TCNCNTYAAARGAYTAYGARGTNAAR 42
Db      75378 TCTCAATACAAAGACAATGAGTCAAG 75352

RESULT 44
US-09-949-002-802/c
; Sequence 802, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 802
; LENGTH: 209632
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(209632)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-802

Query Match      55.4%; Score 18.4; DB 3; Length 209632;
Best Local Similarity 59.3%; Pred. No. 9.2e+02;
Matches 16; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy      16 TCNCNTYAAARGAYTAYGARGTNAAR 42
Db      75378 TCTCAATACAAAGACAATGAGTCAAG 75352

RESULT 45
US-09-949-016-13358/c
; Sequence 13358, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13358
; LENGTH: 260247
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13358

Query Match      55.4%; Score 18.4; DB 3; Length 260247;
Best Local Similarity 51.4%; Pred. No. 9.7e+02;
Matches 19; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy      5 AYGNTGGCNCNTCCNTAAYAARGAYTAYGARGTNA 41
Db      146594 ATGCATAGGCAGACAATACAAAGACCTTGAGGTCAA 146558

RESULT 46
US-08-991-789A-60/c
; Sequence 60, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
;              Smith, John M.
;              Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-08-991-789A-60

Query Match      54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.8%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy      5 AYGNTGGCNCNTCCNTAAYAARGAYTAYGARGTNAAR 42
Db      134 ATGTGTGGCATTCACCATAAATACTCTAAGAGGTTAAA 97

RESULT 47
US-09-062-451-60/c
; Sequence 60, Application US/09062451
```



```
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-60

Query Match 54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 5 AYCNTGGCCNTCCNTAARGAYTAYGARTNAAR 42
Db 134 ATGTGGCATTGACCATTAATACTCTAAGAGGTTAAA 97

RESULT 48
US-09-598-326-60/c
; Sequence 60, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
```

```
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-598-326-60

Query Match 54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 5 AYCNTGGCCNTCCNTAARGAYTAYGARTNAAR 42
Db 134 ATGTGGCATTGACCATTAATACTCTAAGAGGTTAAA 97

RESULT 49
US-09-289-198-60/c
; Sequence 60, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C5
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: US 08/838,762
; EARLIER FILING DATE: 1997-04-09
; EARLIER APPLICATION NUMBER: PCT/US97/00485
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: US 08/700,014
; EARLIER FILING DATE: 1996-08-20
; EARLIER APPLICATION NUMBER: US 08/585,392
; EARLIER FILING DATE: 1996-01-01
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-289-198-60

Query Match 54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 5 AYCNTGGCCNTCCNTAARGAYTAYGARTNAAR 42
Db 134 ATGTGGCATTGACCATTAATACTCTAAGAGGTTAAA 97

RESULT 50
US-09-429-755-60/c
; Sequence 60, Application US/09429755A
```

; Patent No. 6656480
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Mishner, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-429-755-60

Query Match 54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 5 AYCNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 134 ATGTGGCATTGACCATTAATACTCTAAGAGGTAA 97

RESULT 51

US-09-699-295-60/c

; Sequence 60, Application US/09699295

; Patent No. 6828431

; GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.

; APPLICANT: Reed, Steven G.

; APPLICANT: Smith, John M.

; APPLICANT: Mishner, Linda E.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TREATMENT AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.419C10

; CURRENT APPLICATION NUMBER: US/09/699,295

; CURRENT FILING DATE: 2000-10-26

; NUMBER OF SEQ ID NOS: 326

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 60

; LENGTH: 171

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-699-295-60

Query Match 54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 5 AYCNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 134 ATGTGGCATTGACCATTAATACTCTAAGAGGTAA 97

RESULT 52

US-09-534-825A-60/c

; Sequence 60, Application US/09534825A

; Patent No. 6861506

; GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.

; APPLICANT: Smith, John M.

; APPLICANT: Reed, Steven G.
; APPLICANT: Mishner, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C7
; CURRENT APPLICATION NUMBER: US/09/534,825A
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 317
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-534-825A-60

Query Match 54.8%; Score 18.2; DB 3; Length 171;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 5 AYCNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 134 ATGTGGCATTGACCATTAATACTCTAAGAGGTAA 97

RESULT 53

US-09-949-016-107890

; Sequence 107890, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 107890

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-107890

Query Match 54.8%; Score 18.2; DB 3; Length 601;
Best Local Similarity 48.8%; Pred. No. 2e+02;
Matches 20; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 100 CACATTCATCCCATCCCTTAGAAAAACAATTTACTAAAG 140

RESULT 54

US-09-902-540-1562

; Sequence 1562, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

```
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1562
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1562

Query Match          54.8%; Score 18.2; DB 3; Length 669;
Best Local Similarity 48.8%; Pred. No. 2e+02;
Matches 20; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 2 CMCAYGCTGGCCNTCNCNTAAYAARGAYTAYGARGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 CACATGATTGGCTTCTAAATGTAAAGAAGATAAGACAAA 345

RESULT 55
US-09-328-352-13269
; Sequence 1269, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1269
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1269

Query Match          54.8%; Score 18.2; DB 3; Length 864;
Best Local Similarity 48.8%; Pred. No. 2.2e+02;
Matches 20; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYGCTGGCCNTCNCNTAAYAARGAYTAYGARGTNA 41
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 772 GCCATGATTGGTGCTACTGATTAACGCAAGTGA 812

RESULT 56
US-09-328-352-1336/c
; Sequence 1336, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1336
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1336

Query Match          54.8%; Score 18.2; DB 3; Length 1101;
Best Local Similarity 58.6%; Pred. No. 2.4e+02;
Matches 17; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 13 CCNTCNCNTAAYAARGAYTAYGARGTNA 41
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 676 CCGTCTCGATATCAATATATATGAAGTGAA 648

RESULT 57
US-09-134-000C-2009/c
; Sequence 2009, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2009
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2009

Query Match          54.8%; Score 18.2; DB 3; Length 1335;
Best Local Similarity 58.6%; Pred. No. 2.5e+02;
Matches 17; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 13 CCNTCNCNTAAYAARGAYTAYGARGTNA 41
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 887 CCTCGCCAAATAAGGATAATGAAGATA 859

RESULT 58
US-09-949-016-15870
; Sequence 15870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15870
; LENGTH: 45418
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15870

Query Match          54.8%; Score 18.2; DB 3; Length 45418;
Best Local Similarity 48.8%; Pred. No. 7.4e+02;
Matches 20; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYGCTGGCCNTCNCNTAAYAARGAYTAYGARGTNA 41
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24498 GCTCATGCTTCGTAACTCTCTGTGAATATATAAGGAAA 24538

RESULT 59
US-09-949-016-15494/c
; Sequence 15494, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15494
/ LENGTH: 183770
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(183770)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15494

Query Match          54.8%; Score 18.2; DB 3; Length 183770;
Best Local Similarity 48.8%; Pred. No. 1.1e+03;
Matches 17; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 12 GCNTGNCNTAARGAYTAYGARGTNAAR 42
Db 88582 GCCCTCCATCTACCAAGATTATGAGGGGAAG 88552

RESULT 60
US-09-949-016-15868/c
/ Sequence 15868, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15868
/ LENGTH: 374159
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-15868

Query Match          54.8%; Score 18.2; DB 3; Length 374159;
Best Local Similarity 48.8%; Pred. No. 1.2e+03;
Matches 20; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAARGAYTAYGARGTNAAR 41
Db 162195 GCCCACTATTACCTTCTCTTATAAATGTGATGCAGAAA 162155

RESULT 61
US-09-949-016-14720/c
/ Sequence 14720, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
```

```
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14720
/ LENGTH: 390890
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(390890)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14720

Query Match          54.8%; Score 18.2; DB 3; Length 390890;
Best Local Similarity 48.8%; Pred. No. 1.3e+03;
Matches 20; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 2 CNCAYGCGTGGCCNTCNCNTAARGAYTAYGARGTNAAR 42
Db 226842 CACATTTCATCCCATCCCTTAGAAAAACAATTACTAAAG 226802

RESULT 62
US-09-533-559-1061/c
/ Sequence 1061, Application US/09533559
/ Patent No. 6902887
/ GENERAL INFORMATION:
/ APPLICANT: Randy M. Berka
/ APPLICANT: Michael W. Rey
/ APPLICANT: Jeffrey R. Shuster
/ APPLICANT: Sakari Kauppinen
/ APPLICANT: Ib Groth Clausen
/ APPLICANT: Peter Bjarke Olsen
/ TITLE OF INVENTION: Methods For Monitoring Multiple Gene
/ FILE REFERENCE: S849.200-US
/ CURRENT APPLICATION NUMBER: US/09/533,559
/ CURRENT FILING DATE: 2000-03-22
/ EARLIER APPLICATION NUMBER: 09/273,623
/ EARLIER FILING DATE: 1999-03-22
/ NUMBER OF SEQ ID NOS: 7860
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1061
/ LENGTH: 541
/ TYPE: DNA
/ ORGANISM: Fusarium venenatum
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(541)
/ OTHER INFORMATION: n = A,T,C or G
US-09-533-559-1061

Query Match          54.2%; Score 18; DB 3; Length 541;
Best Local Similarity 55.9%; Pred. No. 2.4e+02;
Matches 19; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAARGAYTAYG 34
Db 116 GCTCAGCATGCCCCCNCNCNCTGTCGCAAGGACTTGG 83

RESULT 63
US-09-949-016-151014
/ Sequence 151014, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 151014

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-151014

Query Match 54.2%; Score 18; DB 3; Length 601;

Best Local Similarity 52.5%; Pred. No. 2.4e+02;

Matches 21; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNAYCGNTGCCNTCCNTAYAAAGAYTAYGAGTNA 40

Db 18 GCACCTGCTGTACATGCACATAAAATGCTCTGAAGTGA 57

RESULT 64

US-09-949-016-193564

; Sequence 193564, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 193564

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-193564

Query Match 54.2%; Score 18; DB 3; Length 601;

Best Local Similarity 64.3%; Pred. No. 2.4e+02;

Matches 18; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 13 CCNTCCNTAYAAAGAYTAYGAGTNA 40

Db 151 CCTTCAACATAAAAGATTATGATGTC 178

RESULT 65

US-08-469-667-8

; Sequence 8, Application US/08469667

; Patent No. 5733748

; GENERAL INFORMATION:

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Rosen, Craig

; TITLE OF INVENTION: Colon Specific Genes and Proteins

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,

; ADDRESSEE: Stewart & Olstein

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: NJ

; COUNTRY: USA

; ZIP: 07068-1739

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,667

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Ferraro, Gregory D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 878 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 2..685

US-08-469-667-8

Query Match 54.2%; Score 18; DB 2; Length 878;

Best Local Similarity 52.9%; Pred. No. 2.7e+02;

Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAYGNTGCCNTCCNTAYAAAGAYTAYGARG 37

Db 710 CATTCTGGCATTACATTTTAAATATGTGG 743

RESULT 66

US-09-224-110-8

; Sequence 8, Application US/09224110

; Patent No. 6337195

; GENERAL INFORMATION:

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Rosen, Craig

; TITLE OF INVENTION: Colon Specific Genes and Proteins

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,

; ADDRESSEE: Stewart & Olstein

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: NJ

; COUNTRY: USA

; ZIP: 07068-1739

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/224,110

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/469,667

; FILING DATE: 06-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Ferraro, Gregory D.

REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
US-09-224-110-8

Query Match 54.2%; Score 18; DB 3; Length 878;
Best Local Similarity 52.9%; Pred. No. 2.7e+02;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CAYCMTGGCCNTCCNTAYARGAYTAYGARG 37
Db 710 CATTCCTGGCATTCCACATTTTAAAAAATTATGTGG 743

RESULT 67
US-09-988-292A-8
Sequence 8, Application US/09988292A
Patent No. 6831152
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,292A
FILING DATE: 19-Nov. 6831152-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-988-292A-8
Query Match 54.2%; Score 18; DB 3; Length 878;

Best Local Similarity 52.9%; Pred. No. 2.7e+02;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
Qy 4 CAYCMTGGCCNTCCNTAYARGAYTAYGARG 37
Db 710 CATTCCTGGCATTCCACATTTTAAAAAATTATGTGG 743

RESULT 68
PCT-US95-07289-8
Sequence 8, Application PC/TUS9507289
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
ADDRESSEE: Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07289
FILING DATE: 06-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-265
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
PCT-US95-07289-8
Query Match 54.2%; Score 18; DB 7; Length 878;
Best Local Similarity 52.9%; Pred. No. 2.7e+02;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
Qy 4 CAYCMTGGCCNTCCNTAYARGAYTAYGARG 37
Db 710 CATTCCTGGCATTCCACATTTTAAAAAATTATGTGG 743

RESULT 69
US-09-016-434-850
Sequence 850, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 850:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT01
CLONE: 608819
US-09-016-434-850

Query Match 54.2%; Score 18; DB 3; Length 1512;
Best Local Similarity 52.9%; Pred. No. 3.2e+02;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAYGNTGCCNTCNCNTAYAAARGAYTAYGARG 37
DB 1400 CATTCCTGGCAITTCACATTTTAAAAATTAATGCG 1433

RESULT 70
US-10-044-205A-3
; Sequence 3, Application US/10044205A
; Patent No. 6984502
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; APPLICANT: BANDARU, Rajasekhar
; TITLE OF INVENTION: 69087, 15821, and 15418, Methods and Compositions of Human Protein
; FILE REFERENCE: Uses Thereof
; FILE REFERENCE: 10147-52U1
; CURRENT APPLICATION NUMBER: US/10/044,205A
; CURRENT FILING DATE: 2002-04-19
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/242,428
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/241,884
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,877
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-044-205A-3

Query Match 54.2%; Score 18; DB 4; Length 1659;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 850:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT01
CLONE: 608819
US-09-016-434-850

Query Match 54.2%; Score 18; DB 3; Length 1512;
Best Local Similarity 52.9%; Pred. No. 3.2e+02;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAYGNTGCCNTCNCNTAYAAARGAYTAYGARG 37
DB 1400 CATTCCTGGCAITTCACATTTTAAAAATTAATGCG 1433

RESULT 70
US-10-044-205A-3
; Sequence 3, Application US/10044205A
; Patent No. 6984502
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; APPLICANT: BANDARU, Rajasekhar
; TITLE OF INVENTION: 69087, 15821, and 15418, Methods and Compositions of Human Protein
; FILE REFERENCE: Uses Thereof
; FILE REFERENCE: 10147-52U1
; CURRENT APPLICATION NUMBER: US/10/044,205A
; CURRENT FILING DATE: 2002-04-19
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/242,428
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/241,884
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,877
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-044-205A-3

Query Match 54.2%; Score 18; DB 4; Length 1659;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;

Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 6 YGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
DB 1155 TGCTGGACGAACACCATTCATCAAGATTACAAGGAAAAG 1191

RESULT 71
US-09-738-894A-1
; Sequence 1, Application US/09738894A
; Patent No. 6331423
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL000636
; CURRENT APPLICATION NUMBER: US/09/738,894A
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Human
US-09-738-894A-1

Query Match 54.2%; Score 18; DB 3; Length 1662;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 6 YGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
DB 1155 TGCTGGACGAACACCATTCATCAAGATTACAAGGAAAAG 1191

RESULT 72
US-09-802-117-1
; Sequence 1, Application US/09802117
; Patent No. 6444456
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6444456el Human G-Coupled Protein Receptor Kinases and Polymu
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0147-USA
; CURRENT APPLICATION NUMBER: US/09/802,117
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/188,449
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-802-117-1

Query Match 54.2%; Score 18; DB 3; Length 1662;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 6 YGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
DB 1155 TGCTGGACGAACACCATTCATCAAGATTACAAGGAAAAG 1191

RESULT 73
US-09-964-469-1
; Sequence 1, Application US/09964469
; Patent No. 6579709
; GENERAL INFORMATION:


```
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636DIV
; CURRENT APPLICATION NUMBER: US/09/964,469
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/208,331
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 09/738,894
; PRIOR FILING DATE: 2000-12-18
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-469-1

Query Match      54.2%; Score 18; DB 3; Length 1662;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      6 YGCNTGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db      1155 TGCTGGACGAACACCATTCACCAAGATTACAGGAAAG 1191
      :||| | | | | | | | | | | | | | | | | | | |
      :||| | | | | | | | | | | | | | | | | | | |

RESULT 74
US-10-217-745-1
; Sequence 1, Application US/10217745
; Patent No. 6838275
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6838275el Human G-Coupled Protein Receptor Kinases and
; TITLE OF INVENTION: Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0147-USA
; CURRENT APPLICATION NUMBER: US/10/217,745
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/802,117
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-217-745-1

Query Match      54.2%; Score 18; DB 3; Length 1662;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      6 YGCNTGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db      1155 TGCTGGACGAACACCATTCACCAAGATTACAGGAAAG 1191
      :||| | | | | | | | | | | | | | | | | | | |
      :||| | | | | | | | | | | | | | | | | | | |

RESULT 75
US-10-425-962-1
; Sequence 1, Application US/10425962
; Patent No. 7029894
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636DIV2
; CURRENT APPLICATION NUMBER: US/10/425,962
; CURRENT FILING DATE: 2003-04-30
```

```
; PRIOR APPLICATION NUMBER: 09/964,469
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 09/738,894
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/208,331
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-425-962-1

Query Match      54.2%; Score 18; DB 5; Length 1662;
Best Local Similarity 48.6%; Pred. No. 3.3e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      6 YGCNTGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db      1155 TGCTGGACGAACACCATTCACCAAGATTACAGGAAAG 1191
      :||| | | | | | | | | | | | | | | | | | | |
      :||| | | | | | | | | | | | | | | | | | | |

RESULT 76
US-09-583-110-560
; Sequence 560, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 560
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-560

Query Match      54.2%; Score 18; DB 3; Length 1953;
Best Local Similarity 52.5%; Pred. No. 3.5e+02;
Matches 21; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY      2 CNCAYCNGTGGCCNTCNCNTAYAAAGAYTAYGAGTNA 41
Db      1208 CAGAAGAATTGCCTGCACGTTACAAGGATAAGGTAGTGA 1247
      :||| | | | | | | | | | | | | | | | | | | |
      :||| | | | | | | | | | | | | | | | | | | |

RESULT 77
US-09-107-433-912
; Sequence 912, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; TITLE OF INVENTION: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; TITLE OF INVENTION: THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
```

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denise
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (761)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 912:
SEQUENCE CHARACTERISTICS:
LENGTH: 1953 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1953
SEQUENCE DESCRIPTION: SEQ ID NO: 912:
US-09-107-433-912

Query Match 54.2%; Score 18; DB 3; Length 1953;
Best Local Similarity 52.5%; Pred. NO. 3.5e+02;
Matches 21; Conservative 4; Mismatches 15; Indels

Qy 2 CNCAYGCNTGGCCNCTCNCNTAYAARGAYTAYGAGTNAA 41
| | | | | | | | | | | | | | |
Db 1208 CAGAAGAATTGCCTGCACGTTACAAGGATTAAGGTAGTGAA 1247

```

RESULT 78
US-09-248-796A-837
, Sequence 837, Application US/09248796A
, Patent No. 6747137
, GENERAL INFORMATION:
, APPLICANT: Keith Weinstock et al
, TITLE OF INVENTION: NUCLEIC ACID AND AMINO
, TITLE OF INVENTION: FOR DIAGNOSTICS AND T
, FILE REFERENCES: 1071967.132
, CURRENT APPLICATION NUMBER: US/09/248,796A
, CURRENT FILING DATE: 1999-02-12
, PRIOR APPLICATION NUMBER: US 60/074,725
, PRIOR FILING DATE: 1998-02-13
, PRIOR APPLICATION NUMBER: US 60/096,409
, PRIOR FILING DATE: 1998-08-13
, NUMBER OF SEQ ID NOS: 28208
, SEQ ID NO 837
, LENGTH: 2148
, TYPE: DNA
, ORGANISM: Candida albicans
US-09-248-796A-837

```

```

.Query Match      54.2%; Score 18; DB 3; Length 2148;
.Best Local Similarity 52.5%; Pred. No. 3.6e+02;
Matches 21; Conservative 4; Mismatches 15; Indels

```

QY 1 GCNCAYGCNTGGCCNTCNCNCNTAYAAARGAYTAYGARTNA 40

Db
769 GCCCAGGCATCGCAATTTCTTTGAAAGATGTCGAGGTCA 808

```

RESULT 79
US-10-044-205A-1
; Sequence 1, Application US/10044205A
; Patent No. 6984502
;
; GENERAL INFORMATION:
;
; APPLICANT: KAFELLER-LIBERMANN, Rosana
; APPLICANT: BANDARU, Rajasekhar
; TITLE OF INVENTION: 69087, 15921, and 15411
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 10147-52U1
; CURRENT APPLICATION NUMBER: US/10/044,205A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/242,428
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/241,884
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,877
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2198
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-044-205A-1

```

Query Match 54.2%; Score 18; DB 4; Length 2198;
Best Local Similarity 48.6%;
Pred. No. 3.6e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

Qy 6 YGNTGGCCNTCNCNTAYAAARGAYTAYGARTNAAR 42
:
1445 TGCTGGACGGAACACCATTCAAAGATTACAAGGAAAAG 1481

Dd

```

RESULT 80
US-09-802-117-5
; Sequence 5, Application US/09802117
; Patent No. 644456
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: NO. 644456el Human G
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0147-USA
; CURRENT APPLICATION NUMBER: US/09/802,117
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/188,449
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2249
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-802-117-5

```

Query Match	54.2%	Score 18;	DB 3;	Length 2249;
Best Local Similarity	48.6%	Pred. No. 3.7e+02;		
Matches 18;	Conservative	7;	Mismatches 12;	Indels 0;
Gaps	0;			

Qy

6 YGCNTGGCCNCTCNCNTAYARGAYTAYGAGTNAAR 42
 :||| | | | | :||| :||| :||| :||| :

pB

1508 TGCTGTACCGAACACCATTCAAAAGATTACAAGGAAAAG 1544

RESULT 81
US-10-217-745-5

; Sequence 5, Application US/10217745
; Patent No. 6838275
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6838275el Human G-Coupled Protein Receptor Kinases and
; TITLE OF INVENTION: Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0147-USA
; CURRENT APPLICATION NUMBER: US/10/217,745
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/802,117
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2249
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-745-5

Query Match 54.2%; Score 18; DB 3; Length 2249;
Best Local Similarity 48.6%; Pred. No. 3.7e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

Qy 6 YGNTGGCCNTCCNTAYAARGAYTAYGAGTNAAR 42
Db 1508 TGCTGGACGAACCACTTCAAGATTACAGGAAG 1544

RESULT 82
US-08-757-541-4/c
; Sequence 4, Application US/08757541
; Patent No. 5766877
; GENERAL INFORMATION:
; APPLICANT: Stark, Kevin Lee
; APPLICANT: Luethy, Roland
; TITLE OF INVENTION: NOVEL AGOUTI-RELATED GENE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DEHAVILLAND DRIVE
; CITY: THOUSAND OAKS
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,541
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: OLESKI, NANCY A
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-402A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-757-541-4

Query Match 54.2%; Score 18; DB 2; Length 2371;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 21; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GCNCAYGCNTGGCCNTCCNTAYAARGAYTAYGAGTNAAR 42
Db 1728 GCACACTCTCTGGCCTTCCCTCTGGTAACAGGATGAGGGCAG 1687

RESULT 83
US-09-033-275-4/c
; Sequence 4, Application US/09033275
; Patent No. 6060589
; GENERAL INFORMATION:
; APPLICANT: Stark, Kevin Lee
; APPLICANT: Luethy, Roland
; TITLE OF INVENTION: NOVEL AGOUTI-RELATED GENE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DEHAVILLAND DRIVE
; CITY: THOUSAND OAKS
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033,275
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OLESKI, NANCY A
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-402A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-033-275-4

Query Match 54.2%; Score 18; DB 3; Length 2371;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 21; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GCNCAYGCNTGGCCNTCCNTAYAARGAYTAYGAGTNAAR 42
Db 1728 GCACACTCTCTGGCCTTCCCTCTGGTAACAGGATGAGGGCAG 1687

RESULT 84
US-09-342-581-4/c
; Sequence 4, Application US/09342581
; Patent No. 6203995
; GENERAL INFORMATION:
; APPLICANT: Stark, Kevin Lee
; APPLICANT: Luethy, Roland
; TITLE OF INVENTION: NOVEL AGOUTI-RELATED GENE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DEHAVILLAND DRIVE
; CITY: THOUSAND OAKS
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

RESULT 91
US-10-001-189-67
; Sequence 67, Application US/10001189
; Patent No. 6962810
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.

```
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 7411
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pBSII-Act5c-ori sequence
US-10-001-189-67

Query Match          54.2%; Score 18; DB 3; Length 7411;
Best Local Similarity 51.4%; Pred. No. 5.3e+02;
Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY      8  CMTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2849  CGTGACCATCACAGCATAAAGATACCGCTGAAG 2883

RESULT 92
US-09-949-016-17104
; Sequence 17104, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17104
; LENGTH: 15566
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17104

Query Match          54.2%; Score 18; DB 3; Length 15566;
Best Local Similarity 56.8%; Pred. No. 6.7e+02;
Matches 21; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY      1  GCNCAYGNTGGCCNTCNCNTAYAAAGAYTAYGARG 37
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1893  GCAGAAGCCTGACCATCCCTAGAGGGCCTGGGAGG 1929

RESULT 93
US-09-738-894A-3
; Sequence 3, Application US/09738894A
; Patent No. 6331423
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636DIV2
```

```
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636
; CURRENT APPLICATION NUMBER: US/09/738,894A
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36651
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-738-894A-3

Query Match          54.2%; Score 18; DB 3; Length 36651;
Best Local Similarity 48.6%; Pred. No. 8.6e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      6  YGCNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
      :||| :||| :||| :||| :||| :||| :||| :|||
Db      31336  TGCTGACGACACACCATTCAAAGATTACAAGGAAAG 31372

RESULT 94
US-09-964-469-3
; Sequence 3, Application US/09964469
; Patent No. 6579709
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636DIV
; CURRENT APPLICATION NUMBER: US/09/964,469
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/208,331
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 09/738,894
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36651
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-964-469-3

Query Match          54.2%; Score 18; DB 3; Length 36651;
Best Local Similarity 48.6%; Pred. No. 8.6e+02;
Matches 18; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      6  YGCNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
      :||| :||| :||| :||| :||| :||| :||| :|||
Db      31336  TGCTGACGACACACCATTCAAAGATTACAAGGAAAG 31372

RESULT 95
US-10-425-962-3
; Sequence 3, Application US/10425962
; Patent No. 7029894
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636DIV2
```

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15890
; LENGTH: 144922
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15890

Query Match 54.2%; Score 18; DB 3; Length 144922;
Best Local Similarity 51.4%; Pred. No. 1.3e+03;
Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 6 YGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNA 40
Db 102181 TGCTAGGCCCTGAACGTATAGGATGATAAAGACA 102215

RESULT 100
US-09-949-016-15980
; Sequence 15980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15980
; LENGTH: 239527
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(239527)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980

Query Match 54.2%; Score 18; DB 3; Length 239527;
Best Local Similarity 52.5%; Pred. No. 1.4e+03;
Matches 21; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 GCNCAYGCNTGCCNTCNCNTAYAAAGAYTAYGARGTNA 40
Db 115769 GCACCTGGCTGTACATGCACATAAATGTCTCTGAAGTGA 115808

Search completed: June 10, 2006, 15:38:55
Job time : 158.15 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:34:22 ; Search time 20.3 Seconds
(without alignments)
262.600 Million cell updates/sec

Title: US-10-600-816-33

Perfect score: 33.2

Sequence: 1 gncaygctggtcctcnc.....ayaargaytaygargtarnaar 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 253354 seqs, 63461778 residues

Total number of hits satisfying chosen parameters: 506708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Published Applications_NA_New.*

- 1: /EMC_Celerra_SID33/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 2: /EMC_Celerra_SID33/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /EMC_Celerra_SID33/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /EMC_Celerra_SID33/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /EMC_Celerra_SID33/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 6: /EMC_Celerra_SID33/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 7: /EMC_Celerra_SID33/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /EMC_Celerra_SID33/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	63.3	1101	7	US-11-217-529-82765 Sequence 82765, A
2	20.4	61.4	2292	7	US-11-217-529-78725 Sequence 78725, A
3	18.8	56.6	1161	7	US-11-217-529-5462 Sequence 5462, A
4	18.8	56.6	2871	7	US-11-217-529-76779 Sequence 76779, A
5	18.4	55.4	1624	6	US-10-953-349-1004 Sequence 1004, A
6	18.2	54.8	2118	7	US-11-217-529-75440 Sequence 75440, A
7	18	54.2	118899	7	US-11-189-279-64 Sequence 64, Appl
8	17.8	53.6	2676	7	US-11-217-529-5102 Sequence 5102, Ap
9	17.6	53.0	818	6	US-10-953-349-38583 Sequence 38683, A
10	17.6	53.0	939	7	US-11-217-529-2422 Sequence 2422, Ap
11	17.6	53.0	54550	7	US-11-318-813-42 Sequence 42, Appl
12	17.4	52.4	955	6	US-10-488-619-3026 Sequence 3026, Ap
13	17.4	52.4	1269	6	US-10-471-571A-5437 Sequence 5437, Ap
14	17.4	52.4	1674	7	US-11-217-529-76209 Sequence 76209, A
15	17.4	52.4	1731	7	US-11-217-529-3913 Sequence 3913, Ap
16	17.4	52.4	1897	6	US-10-505-928-269 Sequence 269, App
17	17.4	52.4	2876	7	US-11-293-697-1526 Sequence 1526, Ap
18	17.2	51.8	972	6	US-10-471-571A-5233 Sequence 523, App
19	17.2	51.8	1121	6	US-10-953-349-23606 Sequence 23606, A
20	17.2	51.8	1377	7	US-11-217-529-78099 Sequence 78099, A
21	17	51.2	1349	6	US-10-953-349-29678 Sequence 29678, A
22	17	51.2	1731	7	US-11-217-529-754 Sequence 754, App
23	17	51.2	1800	7	US-11-217-529-1835 Sequence 1835, Ap
24	17	51.2	2269	7	US-11-293-697-51 Sequence 51, Appl
25	17	51.2	2856	7	US-11-293-697-2015 Sequence 2015, Ap

Sequence 31575, A
Sequence 1736, Ap
Sequence 1, Appli
Sequence 3130, Ap
Sequence 112, App
Sequence 679, App
Sequence 245, App
Sequence 537, App
Sequence 19, Appl
Sequence 5, Appli
Sequence 42, Appl
GENERAL INFORMATI
Sequence 2144, Ap
Sequence 14311, A
Sequence 2143, Ap
Sequence 31441, A
Sequence 25421, A
Sequence 8766, Ap
Sequence 76864, A
Sequence 78513, A
Sequence 13812, A
Sequence 1977, Ap
Sequence 1795, Ap
Sequence 39551, A
Sequence 24567, A
Sequence 35818, A
Sequence 10183, A
Sequence 79698, A
Sequence 2093, Ap
Sequence 771, App
Sequence 76276, A
Sequence 834, App
Sequence 289, App
Sequence 554, App
Sequence 4, Appli
Sequence 208, App
Sequence 1345, Ap
Sequence 25896, A
Sequence 21915, A
Sequence 20753, A
Sequence 75494, A
Sequence 1142, Ap
Sequence 2174, Ap
Sequence 2142, Ap
Sequence 9257, Ap
Sequence 2382, Ap
Sequence 4479, Ap
Sequence 6203, Ap
Sequence 1363, Ap
Sequence 1375, Ap
Sequence 40221, A
Sequence 4801, Ap
Sequence 472, App
Sequence 2852, Ap
Sequence 489, App
Sequence 79889, A
Sequence 3165, Ap
Sequence 6898, Ap
Sequence 35524, A
Sequence 29558, A
Sequence 21144, A
Sequence 21144, A
Sequence 24287, A
Sequence 12909, A
Sequence 71, Appl
Sequence 2, Appli
Sequence 76446, A
Sequence 286, App
Sequence 2992, Ap
Sequence 1298, Ap
Sequence 81691, A
Sequence 1078, Ap
Sequence 79043, A
Sequence 3383, Ap

245	15.2	45.8	2079	7	US-11-217-529-575	Sequence 575, App	C 318	15	45.2	1970	6	US-10-953-349-10697	Sequence 10697, A
246	15.2	45.8	2154	6	US-10-953-349-7367	Sequence 7367, App	319	15	45.2	2013	7	US-11-217-529-1571	Sequence 1571, App
C 247	15.2	45.8	2256	7	US-11-217-529-4066	Sequence 4066, App	320	15	45.2	2075	6	US-10-511-937-454	Sequence 454, App
248	15.2	45.8	2277	7	US-11-233-697-912	Sequence 912, App	C 321	15	45.2	2097	7	US-11-233-697-2195	Sequence 2195, App
C 249	15.2	45.8	2382	7	US-11-233-697-461	Sequence 461, App	C 322	15	45.2	2136	6	US-10-953-349-6908	Sequence 6908, App
250	15.2	45.8	2457	7	US-11-217-529-78765	Sequence 78765, A	323	15	45.2	2226	7	US-11-217-529-77979	Sequence 77979, A
251	15.2	45.8	2462	7	US-11-233-697-2214	Sequence 2214, App	C 324	15	45.2	2250	7	US-11-217-529-3659	Sequence 3659, App
252	15.2	45.8	2479	7	US-11-233-697-2260	Sequence 2260, App	325	15	45.2	2364	7	US-11-217-529-76460	Sequence 76460, A
253	15.2	45.8	2571	7	US-11-217-529-4236	Sequence 4236, App	326	15	45.2	2448	7	US-11-217-529-358	Sequence 358, App
C 254	15.2	45.8	2574	7	US-11-233-697-674	Sequence 674, App	C 327	15	45.2	2472	7	US-11-217-529-1324	Sequence 1324, App
255	15.2	45.8	2660	7	US-11-233-697-2294	Sequence 2294, App	C 328	15	45.2	2473	6	US-10-505-928-264	Sequence 264, App
256	15.2	45.8	2788	7	US-11-233-697-1440	Sequence 1440, App	C 329	15	45.2	2473	6	US-10-511-937-489	Sequence 489, App
C 257	15.2	45.8	2814	7	US-11-217-529-3084	Sequence 3084, App	C 330	15	45.2	2479	7	US-11-233-697-1966	Sequence 1966, App
C 258	15.2	45.8	2850	7	US-11-233-697-2226	Sequence 2226, App	C 331	15	45.2	2545	7	US-11-233-697-1149	Sequence 1149, App
C 259	15.2	45.8	2902	6	US-10-953-349-11826	Sequence 11826, A	C 332	15	45.2	2607	7	US-11-217-529-4263	Sequence 4263, App
C 260	15.2	45.8	2931	7	US-11-217-529-3221	Sequence 3221, App	C 333	15	45.2	2607	7	US-11-233-697-1330	Sequence 1330, App
C 261	15.2	45.8	3432	7	US-11-233-697-1690	Sequence 1690, App	C 334	15	45.2	2819	7	US-11-233-697-760	Sequence 760, App
C 262	15.2	45.8	3531	7	US-11-311-778-23	Sequence 23, App1	C 335	15	45.2	2873	7	US-11-233-697-727	Sequence 727, App
C 263	15.2	45.8	3534	7	US-11-311-778-13	Sequence 13, App1	C 336	15	45.2	2874	7	US-11-217-529-3288	Sequence 3288, App
C 264	15.2	45.8	3663	7	US-11-217-529-1369	Sequence 1369, App	C 337	15	45.2	2898	7	US-11-217-529-156	Sequence 156, App
C 265	15.2	45.8	3732	6	US-11-233-697-826	Sequence 826, App	C 338	15	45.2	2949	7	US-11-233-697-426	Sequence 426, App
C 266	15.2	45.8	3912	6	US-10-518-039-3	Sequence 3, App1	C 339	15	45.2	3032	6	US-10-514-535-7	Sequence 7, App1
C 267	15.2	45.8	4045	6	US-10-511-937-3116	Sequence 3116, App	340	15	45.2	3161	7	US-11-233-697-1992	Sequence 1992, App
C 268	15.2	45.8	4173	6	US-10-518-039-4	Sequence 4, App1	C 341	15	45.2	3198	6	US-10-953-349-8411	Sequence 8411, App
C 269	15.2	45.8	4326	7	US-11-121-154-160	Sequence 160, App	C 342	15	45.2	3270	6	US-10-480-962-24	Sequence 24, App1
C 270	15.2	45.8	6243	6	US-10-518-039-2	Sequence 2, App1	C 343	15	45.2	3354	6	US-11-217-529-75390	Sequence 75390, A
C 271	15.2	45.8	6843	6	US-10-518-039-1	Sequence 1, App1	C 344	15	45.2	3374	6	US-10-514-535-5	Sequence 5, App1
C 272	15.2	45.8	7379	6	US-10-522-356-23	Sequence 23, App1	C 345	15	45.2	3457	7	US-11-233-697-1224	Sequence 1224, App
C 273	15.2	45.8	39536	6	US-10-857-660-27	Sequence 27, App1	C 346	15	45.2	3671	6	US-10-196-749-265	Sequence 265, App
C 274	15	45.2	51	7	US-11-143-642-280	Sequence 280, App	C 347	15	45.2	3733	6	US-10-514-535-3	Sequence 3, App1
C 275	15	45.2	207	7	US-11-217-529-82209	Sequence 82209, A	C 348	15	45.2	3780	6	US-10-509-131-24	Sequence 24, App1
C 276	15	45.2	496	7	US-11-298-560-1	Sequence 1, App1	C 349	15	45.2	3787	6	US-10-514-535-4	Sequence 4, App1
C 277	15	45.2	539	6	US-10-488-619-2573	Sequence 2573, App	C 350	15	45.2	4291	6	US-10-509-131-42	Sequence 42, App1
C 278	15	45.2	563	6	US-11-217-529-4898	Sequence 4898, App	C 351	15	45.2	4626	7	US-11-217-529-79591	Sequence 79591, A
C 279	15	45.2	616	6	US-10-953-349-26873	Sequence 26873, A	C 352	15	45.2	4767	6	US-10-505-928-369	Sequence 369, App
C 280	15	45.2	645	6	US-10-953-349-27697	Sequence 27697, A	C 353	15	45.2	4767	6	US-11-145-307A-150	Sequence 150, App
C 281	15	45.2	684	6	US-10-471-571A-17	Sequence 17, App	C 354	15	45.2	4839	6	US-10-560-723-166	Sequence 166, App
C 282	15	45.2	728	6	US-10-488-619-2229	Sequence 2229, App	C 355	15	45.2	5988	7	US-11-217-529-505	Sequence 505, App
C 283	15	45.2	730	6	US-10-953-349-2935	Sequence 2935, App	C 356	15	45.2	27439	6	US-10-857-260-34	Sequence 34, App1
C 284	15	45.2	739	6	US-10-953-349-11443	Sequence 11443, A	C 357	15	45.2	151830	6	US-10-519-335-37	Sequence 37, App1
C 285	15	45.2	867	7	US-11-298-560-25	Sequence 25, App1	C 358	14.8	44.6	25	7	US-11-217-529-172215	Sequence 172215, A
C 286	15	45.2	914	6	US-10-953-349-6898	Sequence 6898, App	C 359	14.8	44.6	25	7	US-11-217-529-192756	Sequence 192756, A
C 287	15	45.2	921	6	US-10-953-349-6843	Sequence 6843, App	C 360	14.8	44.6	44	6	US-10-514-776-283	Sequence 283, App
C 288	15	45.2	973	6	US-10-953-349-17349	Sequence 17349, A	C 361	14.8	44.6	51	7	US-11-143-642-1387	Sequence 1387, App
C 289	15	45.2	1090	6	US-10-953-349-39412	Sequence 39412, A	C 362	14.8	44.6	192	7	US-11-217-529-166960	Sequence 166960, A
C 290	15	45.2	1099	6	US-10-953-349-26935	Sequence 26935, A	C 363	14.8	44.6	192	7	US-11-217-529-190954	Sequence 190954, A
C 291	15	45.2	1074	7	US-11-217-529-77518	Sequence 77518, A	C 364	14.8	44.6	237	6	US-10-473-173-172	Sequence 172, App
C 292	15	45.2	1090	6	US-10-953-349-3174	Sequence 3174, App	C 365	14.8	44.6	243	7	US-11-217-529-82301	Sequence 82301, A
C 293	15	45.2	1124	6	US-10-953-349-33000	Sequence 33000, A	C 366	14.8	44.6	282	7	US-11-217-529-82301	Sequence 82301, A
C 294	15	45.2	1146	6	US-10-953-349-20608	Sequence 20608, A	C 367	14.8	44.6	372	7	US-11-301-554-536	Sequence 536, App
C 295	15	45.2	1159	6	US-10-953-349-30736	Sequence 30736, A	C 368	14.8	44.6	372	7	US-11-301-554-603	Sequence 603, App
C 296	15	45.2	1251	7	US-11-217-529-77970	Sequence 77970, A	C 369	14.8	44.6	454	7	US-11-301-554-1238	Sequence 1238, App
C 297	15	45.2	1272	6	US-10-953-349-4250	Sequence 4250, App	C 370	14.8	44.6	456	7	US-11-217-529-173259	Sequence 173259, A
C 298	15	45.2	1326	6	US-10-953-349-17594	Sequence 17594, A	C 371	14.8	44.6	522	6	US-10-488-619-1461	Sequence 1461, App
C 299	15	45.2	1347	7	US-11-217-529-3080	Sequence 3080, App	C 372	14.8	44.6	551	6	US-10-953-349-26465	Sequence 26465, A
C 300	15	45.2	1357	6	US-10-953-349-23753	Sequence 23753, A	C 373	14.8	44.6	596	6	US-10-488-619-536	Sequence 536, App
C 301	15	45.2	1403	6	US-10-953-349-19405	Sequence 19405, A	C 374	14.8	44.6	601	1	US-09-949-925-17	Sequence 17, App1
C 302	15	45.2	1414	6	US-10-953-349-3678	Sequence 3678, App	C 375	14.8	44.6	651	7	US-11-217-529-76681	Sequence 76681, A
C 303	15	45.2	1431	6	US-10-953-349-435	Sequence 435, App	C 376	14.8	44.6	702	7	US-11-217-529-81658	Sequence 81658, A
C 304	15	45.2	1466	6	US-10-953-349-6648	Sequence 6648, App	C 377	14.8	44.6	729	6	US-10-953-349-33114	Sequence 33114, A
C 305	15	45.2	1491	6	US-10-953-349-22962	Sequence 22962, A	C 378	14.8	44.6	750	6	US-10-488-619-1179	Sequence 1179, App
C 306	15	45.2	1506	7	US-11-217-529-146	Sequence 146, App	C 379	14.8	44.6	774	6	US-10-471-571A-683	Sequence 683, App
C 307	15	45.2	1596	7	US-11-217-529-80857	Sequence 80857, A	C 380	14.8	44.6	804	7	US-11-217-529-4180	Sequence 4180, App
C 308	15	45.2	1599	6	US-10-953-349-32672	Sequence 32672, A	C 381	14.8	44.6	854	7	US-11-217-529-190872	Sequence 190872, A
C 309	15	45.2	1618	6	US-10-623-808-3	Sequence 3, App1	C 382	14.8	44.6	878	6	US-10-953-349-6367	Sequence 6367, App
C 310	15	45.2	1672	7	US-11-101-316-17	Sequence 17, App1	C 383	14.8	44.6	886	6	US-10-953-349-723	Sequence 723, App
C 311	15	45.2	1716	7	US-11-217-529-79923	Sequence 79923, A	C 384	14.8	44.6	888	7	US-11-217-529-76968	Sequence 76968, A
C 312	15	45.2	1761	6	US-10-953-349-38337	Sequence 38337, A	C 385	14.8	44.6	897	7	US-11-217-529-78256	Sequence 78256, A
C 313	15	45.2	1818	6	US-10-953-349-31933	Sequence 31933, A	C 386	14.8	44.6	956	6	US-10-953-349-4195	Sequence 4195, App
C 314	15	45.2	1845	6	US-10-953-349-33183	Sequence 33183, A	C 387	14.8	44.6	1000	6	US-10-953-349-37594	Sequence 37594, A
C 315	15	45.2	1846	6	US-10-953-349-24199	Sequence 24199, A	C 388	14.8	44.6	1002	7	US-11-233-726-2	Sequence 2, App1
C 316	15	45.2	1896	7	US-11-217-529-2887	Sequence 2887, App	C 389	14.8	44.6	1008	7	US-11-217-529-1836	Sequence 1836, App
C 317	15	45.2	1919	6	US-10-953-349-188	Sequence 188, App	C 390	14.8	44.6	1061	6	US-10-953-349-25230	Sequence 25230, A

C 391	14.8	44.6	1065	7	US-11-217-529-79313	Sequence 79313, A	464	14.8	44.6	4585	6	US-10-505-928-47	Sequence 47, Appl
C 392	14.8	44.6	1098	6	US-10-953-349-26400	Sequence 26400, A	C 465	14.8	44.6	4638	7	US-11-217-529-6212	Sequence 6212, Ap
C 393	14.8	44.6	1143	7	US-11-217-529-77278	Sequence 77278, A	C 466	14.8	44.6	4798	6	US-10-505-928-372	Sequence 372, App
C 394	14.8	44.6	1149	7	US-11-217-529-1338	Sequence 1338, Ap	C 467	14.8	44.6	7268	6	US-10-559-415-189	Sequence 189, App
C 395	14.8	44.6	1173	7	US-11-242-505A-26	Sequence 26, Appl	C 468	14.8	44.6	7430	6	US-10-559-415-1	Sequence 1, Appli
C 396	14.8	44.6	1203	6	US-10-471-571A-651	Sequence 651, App	C 469	14.8	44.6	56580	6	US-10-553-298-1	Sequence 1, Appli
C 397	14.8	44.6	1232	6	US-10-953-349-16959	Sequence 16959, A	C 470	14.8	44.6	128361	6	US-10-505-928-151	Sequence 151, App
C 398	14.8	44.6	1239	6	US-10-953-349-15985	Sequence 15985, A	C 471	14.6	44.0	111	7	US-11-217-529-78759	Sequence 78759, A
C 399	14.8	44.6	1252	7	US-11-140-450-91	Sequence 91, Appl	C 472	14.6	44.0	210	6	US-10-505-928-116	Sequence 116, App
C 400	14.8	44.6	1265	7	US-11-242-505A-25	Sequence 25, Appl	C 473	14.6	44.0	324	6	US-10-473-173-374	Sequence 374, App
C 401	14.8	44.6	1316	6	US-11-242-505A-122	Sequence 122, App	C 474	14.6	44.0	334	6	US-10-471-571A-2063	Sequence 2063, Ap
C 402	14.8	44.6	1325	6	US-10-953-349-7487	Sequence 7487, Ap	C 475	14.6	44.0	357	7	US-11-217-529-79003	Sequence 79003, A
C 403	14.8	44.6	1331	6	US-10-953-349-10757	Sequence 10757, A	C 476	14.6	44.0	360	7	US-11-217-529-79003	Sequence 16, Appl
C 404	14.8	44.6	1346	6	US-10-953-349-13179	Sequence 13179, A	C 477	14.6	44.0	399	6	US-10-488-619-26	Sequence 26, Appl
C 405	14.8	44.6	1357	6	US-10-953-349-2188	Sequence 2188, Ap	C 478	14.6	44.0	423	7	US-11-217-529-79004	Sequence 79004, A
C 406	14.8	44.6	1371	6	US-10-953-349-18720	Sequence 18720, A	C 479	14.6	44.0	426	7	US-11-217-529-174276	Sequence 174276,
C 407	14.8	44.6	1395	7	US-11-217-529-76850	Sequence 76850, A	C 480	14.6	44.0	465	6	US-10-488-619-1631	Sequence 1631, Ap
C 408	14.8	44.6	1404	6	Sequence 3454, Ap	Sequence 3454, Ap	C 481	14.6	44.0	480	7	US-11-217-529-80407	Sequence 80407, A
C 409	14.8	44.6	1410	7	US-10-471-571A-5019	Sequence 5019, Ap	C 482	14.6	44.0	493	6	US-10-560-723-141	Sequence 141, App
C 410	14.8	44.6	1422	6	US-11-217-529-2085	Sequence 2085, Ap	C 483	14.6	44.0	496	6	US-10-488-619-1051	Sequence 1051, App
C 411	14.8	44.6	1425	6	US-10-953-349-30001	Sequence 30001, A	C 484	14.6	44.0	506	6	US-10-953-349-26326	Sequence 26326, A
C 412	14.8	44.6	1458	7	US-11-217-529-4057	Sequence 4057, Ap	C 485	14.6	44.0	525	7	US-11-301-554-346	Sequence 346, App
C 413	14.8	44.6	1514	6	US-10-953-349-17445	Sequence 17445, A	C 486	14.6	44.0	525	7	US-11-301-554-346	Sequence 346, App
C 414	14.8	44.6	1536	7	US-11-217-529-3585	Sequence 3585, Ap	C 487	14.6	44.0	531	6	US-10-471-571A-4531	Sequence 4531, Ap
C 415	14.8	44.6	1540	6	US-10-953-349-23951	Sequence 23951, A	C 488	14.6	44.0	586	6	US-10-953-349-25377	Sequence 25377, A
C 416	14.8	44.6	1553	6	US-10-473-173-115	Sequence 115, App	C 489	14.6	44.0	594	6	US-10-471-571A-1277	Sequence 1277, Ap
C 417	14.8	44.6	1560	6	US-10-953-349-11211	Sequence 11211, A	C 490	14.6	44.0	594	6	US-10-488-619-1656	Sequence 1656, Ap
C 418	14.8	44.6	1567	7	US-11-293-697-1059	Sequence 1059, Ap	C 491	14.6	44.0	603	6	US-10-953-349-28016	Sequence 28016, A
C 419	14.8	44.6	1574	6	US-10-953-349-8516	Sequence 8516, Ap	C 492	14.6	44.0	628	6	US-10-953-349-16136	Sequence 16136, A
C 420	14.8	44.6	1590	6	US-11-217-529-79003	Sequence 79003, A	C 493	14.6	44.0	754	6	US-10-953-349-31094	Sequence 31094, A
C 421	14.8	44.6	1610	6	US-10-953-349-40042	Sequence 40042, A	C 494	14.6	44.0	777	7	US-11-217-529-173451	Sequence 173451,
C 422	14.8	44.6	1700	6	US-10-953-349-121788	Sequence 121788, A	C 495	14.6	44.0	786	7	US-11-217-529-83	Sequence 83, Appl
C 423	14.8	44.6	1727	6	US-10-953-349-12509	Sequence 12509, A	C 496	14.6	44.0	786	7	US-11-217-529-75420	Sequence 75420, A
C 424	14.8	44.6	1728	6	US-10-953-349-22752	Sequence 22752, A	C 497	14.6	44.0	805	7	US-11-217-529-77431	Sequence 77431, A
C 425	14.8	44.6	1749	6	US-10-953-349-24812	Sequence 24812, A	C 498	14.6	44.0	834	6	US-10-953-349-26784	Sequence 26784, A
C 426	14.8	44.6	1758	6	Sequence 75573, A	Sequence 75573, A	C 499	14.6	44.0	849	7	US-11-217-529-76262	Sequence 76262, A
C 427	14.8	44.6	1823	7	US-11-293-697-2080	Sequence 2080, Ap	C 500	14.6	44.0	878	6	US-10-953-349-34165	Sequence 34165, A
C 428	14.8	44.6	1827	7	US-11-203-828-4	Sequence 4, Appli	C 501	14.6	44.0	878	6	US-10-953-349-26778	Sequence 26778, A
C 429	14.8	44.6	1839	7	US-11-217-529-191169	Sequence 191169,	C 502	14.6	44.0	885	6	US-10-471-571A-3451	Sequence 3451, Ap
C 430	14.8	44.6	1908	7	US-11-217-529-609	Sequence 609, App	C 503	14.6	44.0	923	6	US-10-953-349-18891	Sequence 18891, A
C 431	14.8	44.6	1916	7	US-11-236-238-28	Sequence 28, Appl	C 504	14.6	44.0	947	6	US-10-953-349-12552	Sequence 12552, A
C 432	14.8	44.6	2027	7	US-11-293-697-1275	Sequence 1275, Ap	C 505	14.6	44.0	999	6	US-10-471-571A-2423	Sequence 2423, Ap
C 433	14.8	44.6	2057	7	US-11-293-697-1042	Sequence 1042, Ap	C 506	14.6	44.0	1043	6	US-10-953-349-28829	Sequence 28829, A
C 434	14.8	44.6	2168	6	US-11-293-697-2011	Sequence 2011, Ap	C 507	14.6	44.0	1056	6	US-10-953-349-39007	Sequence 39007, A
C 435	14.8	44.6	2206	6	US-10-953-349-16892	Sequence 16892, A	C 508	14.6	44.0	1075	6	US-10-953-349-34079	Sequence 34079, A
C 436	14.8	44.6	2225	6	US-10-505-928-24	Sequence 24, Appl	C 509	14.6	44.0	1079	6	US-10-953-349-39132	Sequence 39132, A
C 437	14.8	44.6	2232	7	US-11-217-529-5909	Sequence 5909, Ap	C 510	14.6	44.0	1095	6	US-10-196-749-231	Sequence 231, App
C 438	14.8	44.6	2240	7	US-11-293-697-1451	Sequence 1451, Ap	C 511	14.6	44.0	1107	6	US-10-953-349-9373	Sequence 9373, Ap
C 439	14.8	44.6	2257	6	US-10-953-349-35594	Sequence 35594, A	C 512	14.6	44.0	1116	7	US-11-217-529-1234	Sequence 1234, Ap
C 440	14.8	44.6	2260	6	US-11-293-697-2317	Sequence 2317, Ap	C 513	14.6	44.0	1127	6	US-10-953-349-9178	Sequence 9178, Ap
C 441	14.8	44.6	2291	7	US-11-293-697-875	Sequence 875, App	C 514	14.6	44.0	1139	6	US-10-953-349-34003	Sequence 34003, A
C 442	14.8	44.6	2299	7	US-11-293-697-2411	Sequence 2411, Ap	C 515	14.6	44.0	1159	6	US-10-953-349-22297	Sequence 22297, A
C 443	14.8	44.6	2364	7	US-11-217-529-2411	Sequence 2411, Ap	C 516	14.6	44.0	1186	6	US-10-953-349-21809	Sequence 21809, A
C 444	14.8	44.6	2406	7	US-11-293-697-239	Sequence 239, App	C 517	14.6	44.0	1203	7	US-11-217-529-82035	Sequence 82035, A
C 445	14.8	44.6	2412	7	US-11-217-529-82371	Sequence 82371, A	C 518	14.6	44.0	1212	7	US-11-217-529-78358	Sequence 78358, A
C 446	14.8	44.6	2482	6	US-10-953-349-10942	Sequence 10942, A	C 519	14.6	44.0	1248	7	US-11-217-529-82401	Sequence 82401, A
C 447	14.8	44.6	2579	6	US-11-293-697-2363	Sequence 2363, Ap	C 520	14.6	44.0	1263	7	US-11-217-529-78073	Sequence 78073, A
C 448	14.8	44.6	2583	7	US-10-473-173-96	Sequence 96, Appl	C 521	14.6	44.0	1305	6	US-10-953-349-3689	Sequence 3689, Ap
C 449	14.8	44.6	2655	6	US-11-217-529-81093	Sequence 81093, A	C 522	14.6	44.0	1318	6	US-10-953-349-17559	Sequence 17559, Ap
C 450	14.8	44.6	2667	7	US-11-217-529-81093	Sequence 81093, A	C 523	14.6	44.0	1334	6	US-10-953-349-9228	Sequence 9228, Ap
C 451	14.8	44.6	2696	6	US-10-473-173-109	Sequence 109, App	C 524	14.6	44.0	1347	7	US-11-217-529-80117	Sequence 80117, A
C 452	14.8	44.6	2922	7	US-11-121-154-68	Sequence 68, Appl	C 525	14.6	44.0	1378	6	US-10-953-349-652	Sequence 652, App
C 453	14.8	44.6	2934	6	US-11-217-529-79433	Sequence 79433, A	C 526	14.6	44.0	1392	7	US-11-217-529-2976	Sequence 2976, Ap
C 454	14.8	44.6	3036	6	US-10-471-571A-641	Sequence 641, App	C 527	14.6	44.0	1398	7	US-11-217-529-726	Sequence 726, App
C 455	14.8	44.6	3102	7	US-11-217-529-82302	Sequence 82302, A	C 528	14.6	44.0	1398	7	US-11-217-529-78184	Sequence 78184, A
C 456	14.8	44.6	3121	7	US-11-293-697-1532	Sequence 1532, Ap	C 529	14.6	44.0	1410	7	US-11-217-529-2085	Sequence 2085, Ap
C 457	14.8	44.6	3319	6	US-10-142-2759-1	Sequence 1, Appli	C 530	14.6	44.0	1421	6	US-10-953-349-22958	Sequence 22958, A
C 458	14.8	44.6	3330	7	US-11-217-529-3415	Sequence 3415, Ap	C 531	14.6	44.0	1426	6	US-10-953-349-21489	Sequence 21489, A
C 459	14.8	44.6	3770	7	US-11-247-437-9	Sequence 9, Appli	C 532	14.6	44.0	1440	6	US-10-953-349-19731	Sequence 19731, A
C 460	14.8	44.6	3822	7	US-11-293-697-877	Sequence 877, App	C 533	14.6	44.0	1446	7	US-11-217-529-5439	Sequence 5439, Ap
C 461	14.8	44.6	4021	7	US-11-293-697-472	Sequence 472, App	C 534	14.6	44.0	1473	7	US-11-217-529-3819	Sequence 3819, Ap
C 462	14.8	44.6	4143	7	US-11-217-529-81170	Sequence 81170, A	C 535	14.6	44.0	1498	6	US-10-953-349-9710	Sequence 9710, Ap
C 463	14.8	44.6	4573	6	US-10-505-928-409	Sequence 409, App	C 536	14.6	44.0	1520	6	US-10-953-349-5515	Sequence 5515, Ap

537	14.6	44.0	1536	7	US-11-217-529-81096	Sequence 81096, A	610	14.4	43.4	482	6	US-10-953-349-26070	Sequence 26070, A
538	14.6	44.0	1539	6	US-10-953-349-13732	Sequence 13732, A	611	14.4	43.4	505	6	US-10-488-619-592	Sequence 592, App
539	14.6	44.0	1541	6	US-10-953-349-35987	Sequence 35987, A	612	14.4	43.4	537	6	US-10-488-619-2695	Sequence 2695, App
540	14.6	44.0	1551	7	US-11-217-529-77104	Sequence 77104, A	c 613	14.4	43.4	541	6	US-10-488-619-2784	Sequence 2784, App
541	14.6	44.0	1573	6	US-10-953-349-33345	Sequence 33345, A	c 614	14.4	43.4	567	6	US-10-488-619-1746	Sequence 1746, App
542	14.6	44.0	1578	6	US-11-217-529-75427	Sequence 75427, A	615	14.4	43.4	577	6	US-10-953-349-13304	Sequence 13304, A
543	14.6	44.0	1603	6	US-10-953-349-40023	Sequence 40023, A	c 616	14.4	43.4	591	7	US-11-217-529-77225	Sequence 77225, A
544	14.6	44.0	1658	6	US-10-953-349-33724	Sequence 33724, A	c 617	14.4	43.4	607	6	US-10-488-619-1390	Sequence 1390, App
545	14.6	44.0	1666	6	US-10-953-349-21211	Sequence 21211, A	c 618	14.4	43.4	607	6	US-10-488-619-2799	Sequence 2799, App
546	14.6	44.0	1701	6	US-11-217-529-5374	Sequence 5374, App	c 619	14.4	43.4	628	7	US-11-177-635-1	Sequence 1, Appl
547	14.6	44.0	1731	6	US-10-953-349-1919	Sequence 1919, App	c 620	14.4	43.4	630	7	US-11-217-529-77404	Sequence 77404, A
548	14.6	44.0	1845	7	US-11-217-529-4525	Sequence 4525, App	621	14.4	43.4	636	7	US-11-217-529-82690	Sequence 82690, A
549	14.6	44.0	1848	7	US-11-217-529-190849	Sequence 190849, A	622	14.4	43.4	696	7	US-11-217-529-82746	Sequence 82746, A
550	14.6	44.0	1849	6	US-11-293-697-1826	Sequence 1826, App	c 623	14.4	43.4	705	6	US-10-953-349-35269	Sequence 35269, A
551	14.6	44.0	1859	6	US-10-953-349-34645	Sequence 34645, A	c 624	14.4	43.4	719	6	US-10-953-349-39711	Sequence 39711, A
552	14.6	44.0	1863	7	US-11-217-529-5185	Sequence 5185, App	625	14.4	43.4	730	6	US-10-488-619-2141	Sequence 2141, App
553	14.6	44.0	1908	6	US-10-953-349-3928	Sequence 3928, App	626	14.4	43.4	736	6	US-10-953-349-28371	Sequence 28371, A
554	14.6	44.0	1911	6	US-10-953-349-32981	Sequence 32981, A	c 627	14.4	43.4	738	6	US-10-471-571A-2939	Sequence 2939, App
555	14.6	44.0	1941	7	US-11-293-697-1961	Sequence 1961, App	c 628	14.4	43.4	741	7	US-11-217-529-81247	Sequence 81247, A
556	14.6	44.0	1943	6	US-10-953-349-20012	Sequence 20012, A	c 629	14.4	43.4	753	7	US-11-217-529-4328	Sequence 4328, App
557	14.6	44.0	1983	7	US-11-217-529-82251	Sequence 82251, A	c 630	14.4	43.4	766	6	US-10-953-349-36494	Sequence 36494, A
558	14.6	44.0	1987	6	US-10-953-349-20741	Sequence 20741, A	c 631	14.4	43.4	771	6	US-10-953-349-32895	Sequence 32895, A
559	14.6	44.0	2008	7	US-11-293-697-75	Sequence 75, Appl	c 632	14.4	43.4	793	6	US-10-953-349-15091	Sequence 15091, A
560	14.6	44.0	2025	7	US-11-217-529-1934	Sequence 1934, App	633	14.4	43.4	840	6	US-10-953-349-9132	Sequence 9132, App
561	14.6	44.0	2031	7	US-11-256-428-8	Sequence 8, Appl	634	14.4	43.4	846	7	US-11-217-529-174377	Sequence 174377, A
562	14.6	44.0	2031	7	US-11-256-428-8	Sequence 8, Appl	635	14.4	43.4	875	6	US-10-953-349-8261	Sequence 8261, App
563	14.6	44.0	2057	7	US-11-293-697-2231	Sequence 2231, App	636	14.4	43.4	885	6	US-10-471-571A-2665	Sequence 2665, App
564	14.6	44.0	2115	7	US-11-217-529-76502	Sequence 76502, A	637	14.4	43.4	926	6	US-10-953-349-38418	Sequence 38418, A
565	14.6	44.0	2117	7	US-11-293-697-2250	Sequence 2250, App	c 638	14.4	43.4	957	7	US-11-217-529-77085	Sequence 77085, A
566	14.6	44.0	2160	7	US-11-217-529-1500	Sequence 1500, App	c 639	14.4	43.4	978	6	US-10-953-349-30985	Sequence 30985, A
567	14.6	44.0	2244	7	US-11-293-697-854	Sequence 854, App	640	14.4	43.4	993	6	US-10-953-349-34335	Sequence 34335, A
568	14.6	44.0	2259	7	US-11-293-697-1478	Sequence 1478, App	641	14.4	43.4	1003	7	US-11-226-605-63	Sequence 63, Appl
569	14.6	44.0	2288	6	US-10-953-349-35869	Sequence 35869, A	642	14.4	43.4	1019	7	US-11-297-134-52	Sequence 52, Appl
570	14.6	44.0	2439	6	US-11-293-697-1431	Sequence 1431, App	643	14.4	43.4	1050	6	US-10-953-349-8754	Sequence 8754, App
571	14.6	44.0	2459	6	US-10-953-349-11593	Sequence 11593, A	644	14.4	43.4	1054	6	US-10-953-349-34248	Sequence 34248, A
572	14.6	44.0	2471	7	US-11-293-697-503	Sequence 503, App	645	14.4	43.4	1076	6	US-10-953-349-632	Sequence 632, App
573	14.6	44.0	2552	7	US-11-293-697-390	Sequence 390, App	646	14.4	43.4	1094	6	US-10-953-349-35709	Sequence 35709, A
574	14.6	44.0	2621	7	US-11-293-697-42	Sequence 42, Appl	647	14.4	43.4	1095	6	US-11-217-529-2532	Sequence 2532, App
575	14.6	44.0	2625	7	US-11-217-529-3881	Sequence 3881, App	648	14.4	43.4	1096	6	US-10-953-349-28493	Sequence 28493, A
576	14.6	44.0	2650	7	US-11-293-697-749	Sequence 749, App	649	14.4	43.4	1103	6	US-10-953-349-23955	Sequence 23955, A
577	14.6	44.0	2661	7	US-11-217-529-3021	Sequence 3021, App	c 650	14.4	43.4	1132	6	US-10-953-349-22696	Sequence 22696, A
578	14.6	44.0	2703	7	US-11-293-697-518	Sequence 518, App	651	14.4	43.4	1143	7	US-11-217-529-80824	Sequence 80824, A
579	14.6	44.0	2751	7	US-11-293-697-815	Sequence 815, App	652	14.4	43.4	1185	6	US-10-953-349-33031	Sequence 33031, A
580	14.6	44.0	2937	7	US-11-217-529-77820	Sequence 77820, A	653	14.4	43.4	1186	6	US-11-217-529-190853	Sequence 190853, A
581	14.6	44.0	3069	7	US-11-217-529-76812	Sequence 76812, A	654	14.4	43.4	1238	6	US-10-953-349-19637	Sequence 19637, A
582	14.6	44.0	3102	7	US-11-217-529-5561	Sequence 5561, App	655	14.4	43.4	1242	6	US-10-511-937-433	Sequence 433, App
583	14.6	44.0	3153	7	US-11-217-529-3035	Sequence 3035, App	656	14.4	43.4	1267	6	US-10-525-647-17	Sequence 17, Appl
584	14.6	44.0	3265	7	US-11-293-697-710	Sequence 710, App	657	14.4	43.4	1273	6	US-10-953-349-31969	Sequence 31969, A
585	14.6	44.0	3325	7	US-11-293-697-1050	Sequence 1050, App	c 658	14.4	43.4	1301	6	US-10-953-349-12951	Sequence 12951, A
586	14.6	44.0	3339	7	US-11-217-529-5429	Sequence 5429, App	659	14.4	43.4	1305	6	US-10-953-349-23308	Sequence 23308, A
587	14.6	44.0	3383	7	US-11-258-360-7	Sequence 7, Appl	c 660	14.4	43.4	1324	6	US-10-953-349-14726	Sequence 14726, A
588	14.6	44.0	3655	6	US-10-511-937-346	Sequence 346, App	c 661	14.4	43.4	1329	7	US-11-217-529-2939	Sequence 2939, App
589	14.6	44.0	3680	6	US-10-196-749-173	Sequence 173, App	662	14.4	43.4	1346	6	US-10-953-349-34301	Sequence 34301, A
590	14.6	44.0	3692	7	US-11-293-697-666	Sequence 666, App	c 663	14.4	43.4	1369	6	US-10-953-349-4692	Sequence 4692, App
591	14.6	44.0	3875	6	US-10-560-723-140	Sequence 140, App	c 664	14.4	43.4	1386	6	US-10-953-349-27359	Sequence 27359, A
592	14.6	44.0	3939	7	US-11-242-111-4	Sequence 4, Appl	665	14.4	43.4	1398	7	US-11-217-529-640	Sequence 640, App
593	14.6	44.0	4230	7	US-11-217-529-4374	Sequence 4374, App	c 666	14.4	43.4	1417	6	US-10-953-349-16499	Sequence 16499, A
594	14.6	44.0	4416	7	US-11-217-529-44	Sequence 44, Appl	c 667	14.4	43.4	1428	6	US-10-953-349-15517	Sequence 15517, A
595	14.6	44.0	4624	7	US-11-223-945-44	Sequence 44, Appl	c 668	14.4	43.4	1449	6	US-10-953-349-11641	Sequence 11641, A
596	14.6	44.0	5424	6	US-10-505-928-496	Sequence 496, App	669	14.4	43.4	1449	6	US-10-953-349-12921	Sequence 12921, A
597	14.6	44.0	8019	7	US-11-217-529-76544	Sequence 76544, A	c 670	14.4	43.4	1461	6	US-10-953-349-23145	Sequence 23145, A
598	14.6	44.0	8448	7	US-11-145-307A-166	Sequence 166, App	c 671	14.4	43.4	1478	1	US-09-949-925-56	Sequence 56, Appl
599	14.6	44.0	26850	7	US-11-318-813-41	Sequence 41, Appl	c 672	14.4	43.4	1482	7	US-11-217-529-5605	Sequence 5605, App
600	14.4	43.4	192	7	US-11-217-529-79787	Sequence 79787, A	c 673	14.4	43.4	1506	7	US-11-217-529-284	Sequence 284, App
601	14.4	43.4	357	7	US-11-291-140-46	Sequence 46, Appl	674	14.4	43.4	1509	7	US-11-217-529-191130	Sequence 191130, A
602	14.4	43.4	375	6	US-10-471-571A-5211	Sequence 5211, App	c 675	14.4	43.4	1518	6	US-10-953-349-14536	Sequence 14536, A
603	14.4	43.4	397	7	US-11-301-554-540	Sequence 540, App	c 676	14.4	43.4	1523	6	US-10-953-349-24645	Sequence 24645, A
604	14.4	43.4	423	7	US-11-321-421-271	Sequence 271, App	c 677	14.4	43.4	1542	6	US-10-953-349-16279	Sequence 16279, A
605	14.4	43.4	425	7	US-11-321-421-13	Sequence 13, Appl	c 678	14.4	43.4	1551	6	US-10-471-571A-3561	Sequence 3561, App
606	14.4	43.4	441	6	US-10-488-619-2694	Sequence 2694, App	679	14.4	43.4	1588	7	US-10-953-349-28680	Sequence 28680, A
607	14.4	43.4	451	6	US-10-488-619-401	Sequence 401, App	c 680	14.4	43.4	1599	7	US-11-217-529-78143	Sequence 78143, A
608	14.4	43.4	465	7	US-11-217-529-4864	Sequence 4864, App	c 681	14.4	43.4	1613	6	US-10-953-349-12521	Sequence 12521, A
609	14.4	43.4	480	7	US-11-248-986-1	Sequence 1, Appl	682	14.4	43.4	1613	6	US-10-953-349-26183	Sequence 26183, A


```
C 975 14.2 42.8 768 7 US-11-266-747-2259 Sequence 2259, Ap
C 976 14.2 42.8 768 7 US-11-266-747-2261 Sequence 2261, Ap
C 977 14.2 42.8 768 7 US-11-266-747-2263 Sequence 2263, Ap
C 978 14.2 42.8 768 7 US-11-266-747-2265 Sequence 2265, Ap
C 979 14.2 42.8 768 7 US-11-266-747-2267 Sequence 2267, Ap
C 980 14.2 42.8 768 7 US-11-266-747-2269 Sequence 2269, Ap
C 981 14.2 42.8 768 7 US-11-266-747-2271 Sequence 2271, Ap
C 982 14.2 42.8 768 7 US-11-266-747-2273 Sequence 2273, Ap
C 983 14.2 42.8 768 7 US-11-266-747-2275 Sequence 2275, Ap
C 984 14.2 42.8 768 7 US-11-266-747-2277 Sequence 2277, Ap
C 985 14.2 42.8 768 7 US-11-266-747-2279 Sequence 2279, Ap
C 986 14.2 42.8 768 7 US-11-266-747-2281 Sequence 2281, Ap
C 987 14.2 42.8 768 7 US-11-266-747-2283 Sequence 2283, Ap
C 988 14.2 42.8 768 7 US-11-266-747-2285 Sequence 2285, Ap
C 989 14.2 42.8 768 7 US-11-266-747-2287 Sequence 2287, Ap
C 990 14.2 42.8 768 7 US-11-266-747-2289 Sequence 2289, Ap
C 991 14.2 42.8 768 7 US-11-266-747-2291 Sequence 2291, Ap
C 992 14.2 42.8 768 7 US-11-266-747-2293 Sequence 2293, Ap
C 993 14.2 42.8 768 7 US-11-266-747-2295 Sequence 2295, Ap
C 994 14.2 42.8 768 7 US-11-266-747-2299 Sequence 2299, Ap
C 995 14.2 42.8 768 7 US-11-266-747-2301 Sequence 2301, Ap
C 996 14.2 42.8 768 7 US-11-266-747-2303 Sequence 2303, Ap
C 997 14.2 42.8 768 7 US-11-266-747-2305 Sequence 2305, Ap
C 998 14.2 42.8 768 7 US-11-266-747-2307 Sequence 2307, Ap
C 999 14.2 42.8 768 7 US-11-266-747-2309 Sequence 2309, Ap
C1000 14.2 42.8 768 7 US-11-266-747-2315 Sequence 2315, Ap
```

ALIGNMENTS

```
RESULT 1
US-11-217-529-82765
; Sequence 82765, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 82765
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82765
```

```
Query Match 63.3%; Score 21; DB 7; Length 1101;
Best Local Similarity 55.3%; Pred. No. 0.67;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 4 CAYGNTGGCCNTCCNTAYAAARGAYTAYGAGTNAAR 41
|||:|||||:|||||:|||||:|||||:|||||:
Db 676 CATGCGTGGCCCACTCAACAAACAAAATGAAGTTAA 713
```

```
RESULT 2
US-11-217-529-78725
; Sequence 78725, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
```

```
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 78725
; LENGTH: 2292
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78725
```

```
Query Match 61.4%; Score 20.4; DB 7; Length 2292;
Best Local Similarity 56.8%; Pred. No. 1.6;
Matches 21; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 6 YGNTGGCCNTCCNTAYAAARGAYTAYGAGTNAAR 42
|||:|||||:|||||:|||||:|||||:|||||:
Db 2228 CGCTGGCCATCACCATACATTGGTTAGGAGGAAA 2264
```

```
RESULT 3
US-11-217-529-5462
; Sequence 5462, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 5462
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5462
```

```
Query Match 56.6%; Score 18.8; DB 7; Length 1161;
Best Local Similarity 54.1%; Pred. No. 7.3;
Matches 20; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 5 AYGNTGGCCNTCCNTAYAAARGAYTAYGAGTNAAR 41
|||:|||||:|||||:|||||:|||||:|||||:
Db 692 ATGCTGCTCTACATCGAACAAGAAAACGAAGTCAA 728
```

```
RESULT 4
US-11-217-529-76779
; Sequence 76779, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
```

```
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 76779
; LENGTH: 2871
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76779

Query Match      56.6%; Score 18.8; DB 7; Length 2871;
Best Local Similarity 54.1%; Pred. No. 10;
Matches 20; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 6 YGNTGGCCNTCCNTAYAAARGAYTAYGARGTNAAR 42
Db 2469 TGCATCGCGCTCTCTGAGCAAGATCAAGAAGCAAG 2505

RESULT 5
US-10-953-349-1004
; Sequence 1004, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 1004
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1004

Query Match      55.4%; Score 18.4; DB 6; Length 1624;
Best Local Similarity 54.3%; Pred. No. 13;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 8 CNTGGCCNTCCNTAYAAARGAYTAYGARGTNAAR 42
Db 154 CCTCGCCATCACCGTACTACAGCATGAAGATAAA 188

RESULT 6
US-11-217-529-75440
; Sequence 75440, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 75440
; LENGTH: 2118
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-75440
```

```
Query Match      54.8%; Score 18.2; DB 7; Length 2118;
Best Local Similarity 48.8%; Pred. No. 17;
Matches 20; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 2 CNCAYGCNTGCCNTCCNTAYAAARGAYTAYGARGTNAAR 42
Db 304 CCCTGTTGGCCTTAGTATTCCAAGACTCCGACTTTAAA 344

RESULT 7
US-11-189-279-64/c
; Sequence 64, Application US/11189279
; Publication No. US20060115829A1
; GENERAL INFORMATION:
; APPLICANT: MAO, LI
; APPLICANT: WANG, JIE
; APPLICANT: LUO, WANG
; TITLE OF INVENTION: A METHOD OF TREATING CANCER
; FILE REFERENCE: UTXC:875US
; CURRENT APPLICATION NUMBER: US/11/189,279
; CURRENT FILING DATE: 2005-07-26
; PRIOR APPLICATION NUMBER: 60/598,554
; PRIOR FILING DATE: 2004-08-03
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 64
; LENGTH: 11899
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-189-279-64

Query Match      54.2%; Score 18; DB 7; Length 118899;
Best Local Similarity 50.0%; Pred. No. 81;
Matches 21; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 1 GCNCAYGCNTGGCCNTCCNTAYAAARGAYTAYGARGTNAAR 42
Db 11890 GGCAGGCCAGGCATTCACTTTATATGTTCTGCAGAGAAA 11849
```

```
RESULT 8
US-11-217-529-5102
; Sequence 5102, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 5102
; LENGTH: 2676
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5102

Query Match      53.6%; Score 17.8; DB 7; Length 2676;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 19; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 5 AYCNTGGCCNTCCNTAYAAARGAYTAYGARGTNAAR 42
Db 379 ATGGCTGGCCATGTCTCATCAAGAACAACACGCTCAAG 416
```

```
RESULT 9
US-10-953-349-38683
; Sequence 38683, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38683
; LENGTH: 818
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-38683

Query Match          53.0%; Score 17.6; DB 6; Length 818;
Best Local Similarity 47.6%; Pred. No. 24;
Matches 20; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 1 GCNAYGCGTGGCCNTCCNCTAYARGAYTAYGARGTNAAR 42
Db 439 GCCTTCGCGCGCGCGCGAGCTCAAGGAGTAGCAACAGAAG 480

RESULT 10
US-11-217-529-2422
; Sequence 2422, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2422
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2422

Query Match          53.0%; Score 17.6; DB 7; Length 939;
Best Local Similarity 51.3%; Pred. No. 25;
Matches 20; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 4 CAYGCGTGGCCNTCCNCTAYARGAYTAYGARGTNAAR 42
Db 32 CAAAATTGGCTAAGACCCTACAAGATTGCAAAAGTAAAA 70

RESULT 11
US-11-318-813-42/c
; Sequence 42, Application US/11318813
; Publication No. US20060105381A1
; GENERAL INFORMATION:
; APPLICANT: Ellipse Biotherapeutics Corporation
; APPLICANT: Peltekova, Vanya D
; APPLICANT: Siminovitch, Katherine A
; APPLICANT: St George-Hyslop, Peter H
```

```
; APPLICANT: Rubin, Laurence A
; APPLICANT: Peltekova, Vanya D
; APPLICANT: Wintle, Richard F
; TITLE OF INVENTION: POLYMORPHISMS OF THE OCTN1 AND OCTN2 CATION TRANSPORTERS ASSOCIATED WITH INFLAMMATORY BOWEL DISORDERS
; FILE REFERENCE: ELLP-020
; CURRENT APPLICATION NUMBER: US/11/318,813
; CURRENT FILING DATE: 2005-12-27
; PRIOR APPLICATION NUMBER: US/10/327,188
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/362,700
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 60/343,338
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/427,529
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/362,717
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 54550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (351)..(351)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (436)..(436)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (935)..(935)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1725)..(1725)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1874)..(1874)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2031)..(2031)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2152)..(2152)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2297)..(2297)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2747)..(2747)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3477)..(3477)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3535)..(3535)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4035)..(4035)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
```


; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1526
; LENGTH: 2876
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1526

Query Match 52.4%; Score 17.4; DB 7; Length 2876;
Best Local Similarity 58.1%; Pred. No. 45;
Matches 18; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 5 AYCNGTGGCCNTCNCNTAYAAARGAYTAYGA 35
Db 920 ATGATTGGCCACCCTTAAAGAAGACTGCAA 890

RESULT 18

US-10-471-571A-523
; Sequence 523, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 523
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-523

Query Match 51.8%; Score 17.2; DB 6; Length 972;
Best Local Similarity 48.7%; Pred. No. 39;
Matches 19; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 4 CAYGNTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db 877 CATTTCATTACCATCATCATTTAAGAAGCTTAGAAGATAAA 915

RESULT 19

US-10-953-349-23606
; Sequence 23606, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 23606
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-23606

Query Match 51.8%; Score 17.2; DB 6; Length 1121;
Best Local Similarity 55.9%; Pred. No. 41;
Matches 19; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 4 CAYGNTGGCCNTCNCNTAYAAARGAYTAYGARG 37
Db 642 CATGGATGGACTTTGGCGGTACAATGATGATG 675

RESULT 20

US-11-217-529-78099
; Sequence 78099, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 78099
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78099

Query Match 51.8%; Score 17.2; DB 7; Length 1377;
Best Local Similarity 55.9%; Pred. No. 44;
Matches 19; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 4 CAYGNTGGCCNTCNCNTAYAAARGAYTAYGARG 37
Db 1180 CATGACAGGCCAGCTCCATATAGGAAGATAATG 1213

RESULT 21

US-10-953-349-29678/c
; Sequence 29678, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 29678
; LENGTH: 1349
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-29678

Query Match 51.2%; Score 17; DB 6; Length 1349;
Best Local Similarity 56.7%; Pred. No. 54;
Matches 17; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 2 CNCAYGNTGGCCNTCNCNTAYAAARGAYT 31
Db 638 CTCCTGCGAGGTCTTCTCTTATCAGGATT 609

RESULT 22

US-11-217-529-754/c
; Sequence 754, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED


```
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 754
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-754

Query Match          51.2%; Score 17; DB 7; Length 1731;
Best Local Similarity 54.8%; Pred. No. 59;
Matches 17; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      7 GCGTGGCCNTCCNTAYAAARGAYTAYGARG 37
Db      358 GCTTGGCAATCATGTTATAAGCACCGAAG 328

RESULT 23
US-11-217-529-1835/c
; Sequence 1835, Application US/11217529
; Publication No. US2006099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1835
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1835

Query Match          51.2%; Score 17; DB 7; Length 1800;
Best Local Similarity 60.7%; Pred. No. 60;
Matches 17; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      11 GCGCCTCTCTCCCTAAGGATGAGAAGT 38
Db      1087 GCGCTTCTCTCCCTAAGGATGAGAAGT 1060

RESULT 24
US-11-293-697-51/c
; Sequence 51, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
```

```
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-51

Query Match          51.2%; Score 17; DB 7; Length 2269;
Best Local Similarity 54.8%; Pred. No. 65;
Matches 17; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      5 AYGCCNTGCCNTCCNTAYAAARGAYTAYGA 35
Db      402 ATGCCTGGACTTCCAAAGTATAAATATCTGA 372

RESULT 25
US-11-293-697-2015/c
; Sequence 2015, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2015
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-2015

Query Match          51.2%; Score 17; DB 7; Length 2856;
Best Local Similarity 54.8%; Pred. No. 70;
Matches 17; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      7 GCGTGGCCNTCCNTAYAAARGAYTAYGARG 37
Db      2678 GCGTGGTCTTCCACACACACAGACAGTGAAG 2648

RESULT 26
US-10-953-349-31575
; Sequence 31575, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31575
; LENGTH: 3007
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-31575

Query Match          51.2%; Score 17; DB 6; Length 3007;
Best Local Similarity 60.7%; Pred. No. 71;
Matches 17; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      11 GCGCNTCCNTAYAAARGAYTAYGARGT 38
Db      757 GGCCATCAGATATAGGATTATCATGT 784
```

```
RESULT 27
US-11-293-697-1736/c
; Sequence 1736, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; PRIOR FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1736
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1736

Query Match          51.2%; Score 17; DB 7; Length 3008;
Best Local Similarity 54.8%; Pred. No. 71;
Matches 17; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy      7 GCNTGGCCNTCNCNTAYAAAGAYTAYGARG 37
Db      2832 GCGTGGTCTTCCACACACAGACAGTGAAG 2802

RESULT 28
US-10-502-993-1/c
; Sequence 1, Application US/10502993
; Publication No. US20060089320A1
; GENERAL INFORMATION:
; APPLICANT: CANTLEY, Lewis C.
; APPLICANT: LAMIA, Katja A.
; APPLICANT: RAMEH, Lucia
; APPLICANT: KAHN, Barbara
; APPLICANT: PERONI, Odile
; TITLE OF INVENTION: MODULATION OF TYPE IIa PHOSPHONOSITIDE PHOSPHATE KINASE
; FILE REFERENCE: B0662.70052U01
; CURRENT APPLICATION NUMBER: US/10/502,993
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: US 60/353,758
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3743
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (481)..(1728)
US-10-502-993-1

Query Match          51.2%; Score 17; DB 6; Length 3743;
Best Local Similarity 51.3%; Pred. No. 77;
Matches 20; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

Qy      2 CNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNA 40
Db      2480 CGCTTTCCTCCCTCCACCATCCCATGAGGATGAACCTTA 2442

RESULT 29
US-11-217-529-3130
; Sequence 3130, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
```

```
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3130
; LENGTH: 4656
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-3130

Query Match          51.2%; Score 17; DB 7; Length 4656;
Best Local Similarity 51.5%; Pred. No. 83;
Matches 17; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy      4 CAYGCGTGGCCNTCNCNTAYAAAGAYTAYGAR 36
Db      538 CATGAATTGCTTTGTCTTATAAAGACTTTGAG 570

RESULT 30
US-10-473-173-112
; Sequence 112, Application US/10473173
; Publication No. US2006008823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 112
; LENGTH: 5003
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-112

Query Match          51.2%; Score 17; DB 6; Length 5003;
Best Local Similarity 52.6%; Pred. No. 85;
Matches 20; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

Qy      1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGT 38
Db      2214 GAACCTGCCTGCCCTTCTCCCAAAGGACAAAGAGGT 2251

RESULT 31
US-10-953-349-679
; Sequence 679, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 679
; LENGTH: 802
```



```
RESULT 36
US-11-318-813-42
; Sequence 42, Application US/11318813
; Publication No. US20060105381A1
; GENERAL INFORMATION:
; APPLICANT: Ellipsis Biotherapeutics Corporation
; APPLICANT: Peltekova, Vanya D
; APPLICANT: Siminovich, Katherine A
; APPLICANT: St George-Hyslop, Peter H
; APPLICANT: Rubin, Laurence A
; APPLICANT: Peltekova, Vanya D
; APPLICANT: Wintle, Richard F
; TITLE OF INVENTION: POLYMORPHISMS OF THE OCTN1 AND OCTN2 CATION TRANSPORTERS ASSOCIATED WITH INFLAMMATORY BOWEL DISORDERS
; FILE REFERENCE: ELLP-020
; CURRENT APPLICATION NUMBER: US/11/318,813
; CURRENT FILING DATE: 2005-12-27
; PRIOR APPLICATION NUMBER: US/10/327,188
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/362,700
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 60/343,338
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/427,529
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/362,717
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 54550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (351)..(351)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (436)..(436)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (935)..(935)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1725)..(1725)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1874)..(1874)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2031)..(2031)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2152)..(2152)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2297)..(2297)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2747)..(2747)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3477)..(3477)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3535)..(3535)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4035)..(4035)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4084)..(4084)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4087)..(4087)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4321)..(4321)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4374)..(4374)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4394)..(4394)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4429)..(4429)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4480)..(4480)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4482)..(4482)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4704)..(4704)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4799)..(4799)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4800)..(4800)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4801)..(4801)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4802)..(4802)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4803)..(4803)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4805)..(4805)
; OTHER INFORMATION: n can be a or t or g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4806)..(4806)
; OTHER INFORMATION: n can be a or t or g or c
```



```
Query Match      50.0%; Score 16.6; DB 6; Length 644;
Best Local Similarity 52.8%; Pred. No. 65;
Matches 19; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy      6 YGNTGGCCNTCCNTAYAAAGAYTAYGARGTNA 41
Db      330 TGCTATGCCGCCCTATAGAAAGACCGATGACAA 295

RESULT 40
US-10-488-619-2143
; Sequence 2143, Application US/10488619
; Publication No. US2006009578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2143
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2143

Query Match      50.0%; Score 16.6; DB 6; Length 851;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 19; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy      4 CAYCGTGGCCNTCCNTAYAAAGAYTAYGARGTNA 41
Db      657 CATCCCGAGCCCTTCTCTGATAAGTCTTGGAGGATA 694

RESULT 41
US-10-953-349-31441
; Sequence 31441, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31441
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-31441

Query Match      50.0%; Score 16.6; DB 6; Length 1266;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 19; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy      1 GCNCAVCGTGGCCNTCCNTAYAAAGAYTAYGARGT 38
Db      702 GACACGGTGGCCGCCGCTGTGACGCGATACCGGT 739

RESULT 42
US-10-953-349-25421/c
; Sequence 25421, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
```

```
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25421
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-25421

Query Match      50.0%; Score 16.6; DB 6; Length 1312;
Best Local Similarity 55.9%; Pred. No. 83;
Matches 19; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Qy      7 GNTGGCCNTCCNTAYAAAGAYTAYGARGTNA 40
Db      241 GCGATGCCGCCCGGTAGAAAGACCGAGGTGA 208

RESULT 43
US-10-953-349-8766/c
; Sequence 8766, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8766
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8766

Query Match      50.0%; Score 16.6; DB 6; Length 1746;
Best Local Similarity 46.3%; Pred. No. 91;
Matches 19; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Qy      2 CNGCAVCGTGGCCNTCCNTAYAAAGAYTAYGARGTNA 42
Db      470 CAGACACTTGTCTCTACCAACACGCGATCTGGAGTGA 430

RESULT 44
US-11-217-529-76864
; Sequence 76864, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76864
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76864
```



```
Query Match          49.4%; Score 16.4; DB 6; Length 1087;
Best Local Similarity 47.6%; Pred. No. 96;
Matches 20; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GCNAYGNTGCCNTCCNTAYAAARGAYTAYGARGTNAAR 42
Db 320 GATCAAGGGTGAATCTTCTCTTACGGAGAATATCAACTGCAG 279

RESULT 50
US-10-953-349-24567/c
; Sequence 24567, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24567
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-24567

Query Match          49.4%; Score 16.4; DB 6; Length 1228;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 20; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

Qy 2 CNCAYGNTGCCNTCCNTAYAAARGAYTAYGARGTNA 41
Db 1038 CTATGCTTGTCCATCACTTAACCAATTATATTGAA 999

RESULT 51
US-10-953-349-35818/c
; Sequence 35818, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35818
; LENGTH: 1406
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (534)..(534)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (659)..(659)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-35818

Query Match          49.4%; Score 16.4; DB 6; Length 1406;
Best Local Similarity 54.8%; Pred. No. 1.1e+02;
Matches 17; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 2 CNCAYGNTGCCNTCCNTAYAAARGAYTA 32
Db 941 CGCATGATGCATCATGCTATAAATACGA 911
```

```
RESULT 52
US-10-953-349-10383
; Sequence 10383, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10383
; LENGTH: 1462
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10383

Query Match          49.4%; Score 16.4; DB 6; Length 1462;
Best Local Similarity 48.6%; Pred. No. 1.1e+02;
Matches 17; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 6 YGNTGGCCNTCCNTAYAAARGAYTAYGARGTNA 40
Db 1216 CGCTTGGTCTCTCCATGTTAGAACTTTGAAGAGA 1250

RESULT 53
US-11-217-529-79698/c
; Sequence 79698, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79698
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79698

Query Match          49.4%; Score 16.4; DB 7; Length 1566;
Best Local Similarity 60.7%; Pred. No. 1.1e+02;
Matches 17; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 8 CNTGGCCNTCCNTAYAAARGAYTAYGA 35
Db 1404 CATTTCTTCACCACAAAAGATTACGA 1377

RESULT 54
US-11-293-697-2093
; Sequence 2093, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
```

```
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2093
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-2093

Query Match          49.4%; Score 16.4; DB 7; Length 1854;
Best Local Similarity 47.6%; Pred. No. 1.2e+02;
Matches 20; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGCCNCCNCTAYARGAYTAYGARGTNAAR 42
Db 1460 GCTCAGGCTGGACCCCGTGACAGGCTGCTGAGGGTAAG 1501

RESULT 55
US-10-505-928-771
; Sequence 771, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/33178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 771
; LENGTH: 2180
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-771

Query Match          49.4%; Score 16.4; DB 6; Length 2180;
Best Local Similarity 47.6%; Pred. No. 1.2e+02;
Matches 20; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGCCNCCNCTAYARGAYTAYGARGTNAAR 42
Db 1724 GCTCAGGCTGGACCCCGTGACAGGCTGCTGAGGGTAAG 1765

RESULT 56
US-11-217-529-76276/c
; Sequence 76276, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76276
; LENGTH: 2235
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76276

Query Match          49.4%; Score 16.4; DB 7; Length 2235;
Best Local Similarity 47.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGCCNCCNCTAYARGAYTAYGARGTNAAR 42
Db 1079 GCTCAGGCTGGACCCCGTGACAGGCTGCTGAGGGTAAG 1120

RESULT 59
US-10-505-928-554/c
; Sequence 554, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

Best Local Similarity 56.7%; Pred. No. 1.2e+02;
Matches 17; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 11 GGCCNTCCNCTAYARGAYTAYGARGTNA 40
Db 320 GGCCATCTCTTGCAGAAAACATGAAGTGA 291

RESULT 57
US-11-293-697-834
; Sequence 834, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 834
; LENGTH: 2426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-834

Query Match          49.4%; Score 16.4; DB 7; Length 2426;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGCCNCCNCTAYARGAYTAYGARGTNA 40
Db 1125 GCATGTGCTCTCTTCTAACTGGAAGACTATGACGACA 1164

RESULT 58
US-11-293-697-289
; Sequence 289, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 289
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-289

Query Match          49.4%; Score 16.4; DB 7; Length 2700;
Best Local Similarity 47.6%; Pred. No. 1.3e+02;
Matches 20; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGCCNCCNCTAYARGAYTAYGARGTNAAR 42
Db 1079 GCTCAGGCTGGACCCCGTGACAGGCTGCTGAGGGTAAG 1120

RESULT 59
US-10-505-928-554/c
; Sequence 554, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
```


US-11-217-529-4479

Query Match 48.8%; Score 16.2; DB 7; Length 2163;
Best Local Similarity 47.4%; Pred. NO. 1.5e+02;
Matches 18; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 5 AYGCGTGGCCNTCNCNTAYAAARGAYTAYGAGTNAAR 42
||: ||| | ||| | : ||| : ||| : ||| :
Db 1247 ATGTTTGTGCGTTTCCCATATGCAGTCTTGGAAAGTTAAG 1284

RESULT 73

```

US-11-217-529-6203
; Sequence 6203, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIRO
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6203
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-6203

```

Qy 2 CNCAYGNTGGCNTCNCNTAYAARGAYTAYGA 35
|||: ||| ||| :||| |||:
Db 640 CACACGGAGCTCTTTTCGTTCAAAGAATATAA 673

RESULT 74

```

US-11-217-529-1363
; Sequence 1363, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1363
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1363

```

QY 1 GCNCAYGCNTGGCCNTCNCNTAAYAARGAYTAYGAR 36
||| ||| : | | | | | | | | | | : | | : | :

Pb 907 GCACACACTTCGCGCATCGATGGAYTAACGACCATCAA 942

RESULT 75

```

US-11-293-697-1375/c
; Sequence 1375, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1375
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1375

```

QY
2 CNCA YGNTGGCNCNTCCNTA YAARGAYTAYGARGTNAAR 42
| | | | | | | | : : : : : :
Dd 1860 CTCATGCATGACTTTCTCGCTATCAGCTCCACGGGGTTCAG 1820

RESULT 76

```

US-10-953-349-40221/c
; Sequence 40221, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED D
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 40221
; LENGTH: 2413
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-40221

```

Qy 5 AYGNTGGCCNTCCNTAARGAYTARGTNA 40
|:| | | | | | | | | | | | | | | |
Db 1736 ATGGTGTCCGGTCCATATGTAATGATGTGA 1701

RESULT 77

```

RESUB.
US-10-471-571A-4801
; Sequence 4801, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12

```


Best Local Similarity 57.1%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
Matches 16; Conservative 4; Mismatches 8;

QY 13 CCNTCCNTCAAAAGGATGACTCA 84
DB 57 CCATGTCCTACAAAGGATGACTCA 84

RESULT 81

US-11-217-529-79889/c
; Sequence 79889, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79889
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79889

Query Match 48.2%; Score 16; DB 7; Length 735;
Best Local Similarity 51.4%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 19; Conservative 4; Mismatches 14;

QY 6 YGNTGCCNTCCNTCAAAAGGATGACTCA 42
DB 151 TGCTTGGGATTCGCGTAAAGAGAAAGAAATTCAA 115

RESULT 82

US-10-471-571A-3165/c
; Sequence 3165, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3165
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3165

Query Match 48.2%; Score 16; DB 6; Length 762;
Best Local Similarity 51.4%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 19; Conservative 4; Mismatches 14;

QY 4 CAYGNTGCCNTCCNTCAAAAGGATGACTCA 40
DB 151 CAGGGTGGCTGTACTATAAAAGATATTTCGTTGA 115

RESULT 83

US-10-953-349-6898/c

; Sequence 6898, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6898
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-6898

Query Match 48.2%; Score 16; DB 6; Length 914;
Best Local Similarity 45.2%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 19; Conservative 6; Mismatches 17;

QY 1 GCNCAYGNTGCCNTCCNTCAAAAGGATGACTCA 42
DB 728 GCTCATGCGTGGCGCAACCCCTTTAAAGCCGAGTCAAG 687

RESULT 84

US-10-953-349-35524
; Sequence 35524, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35524
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-35524

Query Match 48.2%; Score 16; DB 6; Length 1227;
Best Local Similarity 45.2%; Pred. No. 1.5e+02; Indels 0; Gaps 0;
Matches 19; Conservative 6; Mismatches 17;

QY 1 GCNCAYGNTGCCNTCCNTCAAAAGGATGACTCA 42
DB 251 GCTTACGCATGGCTGAGCCATAACACAGATGTTGATGACAAG 292

RESULT 85

US-10-953-349-29958
; Sequence 29958, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 29958
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (962)..(962)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1010)..(1010)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1032)..(1032)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1080)..(1080)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1090)..(1090)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1152)..(1152)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1163)..(1163)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1228)..(1228)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1239)..(1239)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-29958

Query Match 48.2%; Score 16; DB 6; Length 1358;
Best Local Similarity 47.6%; Pred. No. 1.6e+02;
Matches 20; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 945 GTTCATGAAGAGCGCGCCGCCCCCAAGGAAACAAGGATAAG 986

RESULT 86
US-10-953-349-21144/c
; Sequence 21144, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21144
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-21144

Query Match 48.2%; Score 16; DB 6; Length 1371;
Best Local Similarity 55.2%; Pred. No. 1.6e+02;
Matches 16; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Qy 10 TGCGCCNTCCNTAYAAAGAYTAYGARGT 38

Db 242 TGGCCTTCATTTCATTAAGAGCTGTGAGGT 214

RESULT 87

US-10-953-349-24287
; Sequence 24287, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24287
; LENGTH: 1499
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-24287

Query Match 48.2%; Score 16; DB 6; Length 1499;
Best Local Similarity 51.4%; Pred. No. 1.7e+02;
Matches 19; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTCCNTCCNTAYAAAGAYTAYGARGT 38
Db 689 CTCACGCGCTTCAGTCACTCTACAACGATTCGGAATT 725

RESULT 88

US-10-953-349-12909
; Sequence 12909, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12909
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (972)..(972)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1573)..(1573)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-12909

Query Match 48.2%; Score 16; DB 6; Length 1584;
Best Local Similarity 47.5%; Pred. No. 1.7e+02;
Matches 19; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTCCNTCCNTAYAAAGAYTAYGARGTNA 41
Db 37 CTCATGATGCCCTCTCTATCTTCTCCCAATGAATGAA 76

RESULT 89

US-09-949-925-71/c
; Sequence 71, Application US/09949925
; Publication No. US20060099575A9
; GENERAL INFORMATION:

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2006, 15:26:46 ; Search time 1608.6 Seconds
(without alignments)
320.826 Million cell updates/sec

Title: US-10-600-816-33
Perfect score: 33.2
Sequence: 1 gncaygcgtggcncnc.....ayaargaytaygargtnaar 42

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :									
Published Applications NA Main:*									
1:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US07_PUBCOMB	.seq:*				
2:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US08_PUBCOMB	.seq:*				
3:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US09A_PUBCOMB	.seq:*				
4:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US09B_PUBCOMB	.seq:*				
5:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US09C_PUBCOMB	.seq:*				
6:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10A_PUBCOMB	.seq:*				
7:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10B_PUBCOMB	.seq:*				
8:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10C_PUBCOMB	.seq:*				
9:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10D_PUBCOMB	.seq:*				
10:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10E_PUBCOMB	.seq:*				
11:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10F_PUBCOMB	.seq:*				
12:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US10G_PUBCOMB	.seq:*				
13:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US11A_PUBCOMB	.seq:*				
14:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US11B_PUBCOMB	.seq:*				
15:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US11C_PUBCOMB	.seq:*				
16:	/EMC	Celerra	SIDS3	/ptodata/2/pubpna/US11D_PUBCOMB	.seq:*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33.2	100.0	42	8	US-10-600-816-33
2	30	90.4	42	8	US-10-600-816-30
3	30	90.4	497	6	US-10-066-543-1937
4	30	90.4	552	6	US-10-066-543-1811
5	30	90.4	620	3	US-09-969-034-2222
6	30	90.4	634	3	US-09-969-034-3393
7	30	90.4	642	7	US-10-125-968-701
8	30	90.4	1071	9	US-10-712-615-134
9	30	90.4	1212	3	US-09-866-050A-249
10	30	90.4	1212	6	US-10-152-661-249
11	30	90.4	1228	7	US-10-313-542-223
12	30	90.4	1400	13	US-11-060-756-2418
13	30	90.4	1400	13	US-11-060-756-6690
14	30	90.4	1460	10	US-10-936-626-64
15	30	90.4	1460	10	US-10-938-061-64
16	30	90.4	1619	7	US-10-224-289-5
17	30	90.4	1619	9	US-10-935-190-43

18	30	90.4	1788	10	US-10-505-486-196	Sequence 196, App
19	30	90.4	2297	9	US-10-775-920-13	Sequence 13, Appl
20	30	90.4	2302	7	US-10-224-289-3	Sequence 3, Appl
21	30	90.4	2302	8	US-10-240-425-405	Sequence 405, App
22	30	90.4	2302	9	US-10-775-920-9	Sequence 9, Appl
23	30	90.4	2302	10	US-10-510-507-2	Sequence 2, Appl
24	30	90.4	2305	9	US-10-775-920-12	Sequence 12, Appl
25	30	90.4	2316	6	US-10-176-847-59	Sequence 59, Appl
26	30	90.4	2316	15	US-11-080-991-59	Sequence 59, Appl
27	30	90.4	2446	9	US-10-775-920-11	Sequence 11, Appl
28	30	90.4	2456	6	US-10-225-567A-453	Sequence 453, App
29	30	90.4	2456	7	US-10-269-909-63	Sequence 63, Appl
30	30	90.4	2456	7	US-10-269-909-64	Sequence 64, Appl
31	30	90.4	2456	7	US-10-295-027-619	Sequence 619, App
32	30	90.4	2456	8	US-10-600-816-2	Sequence 2, Appl
33	30	90.4	2456	8	US-10-600-816-18	Sequence 18, Appl
34	30	90.4	2456	9	US-10-775-920-10	Sequence 10, Appl
35	30	90.4	2456	10	US-10-936-626-40	Sequence 40, Appl
36	30	90.4	2456	10	US-10-938-061-40	Sequence 40, Appl
37	30	90.4	2456	16	US-11-169-041-32	Sequence 32, Appl
38	30	90.4	2593	7	US-10-264-049-834	Sequence 834, App
39	30	90.4	4239	6	US-10-198-846-10424	Sequence 10424, A
40	22.4	67.5	365	8	US-10-424-599-109942	Sequence 109942, A
41	21	63.3	406	4	US-09-925-065A-342206	Sequence 342206, A
42	21	63.3	406	4	US-09-925-065A-342207	Sequence 342207, A
43	21	63.3	406	5	US-09-925-065A-342206	Sequence 342206, A
44	21	63.3	406	5	US-09-925-065A-342207	Sequence 342207, A
45	21	63.3	409	12	US-10-301-480-414987	Sequence 414987, A
46	21	63.3	409	12	US-10-301-480-414988	Sequence 414988, A
47	21	63.3	409	12	US-10-301-480-1028396	Sequence 1028396, A
48	21	63.3	409	12	US-10-301-480-1028397	Sequence 1028397, A
49	21	63.3	552	4	US-09-925-065A-407850	Sequence 407850, A
50	21	63.3	552	5	US-09-925-065A-407850	Sequence 407850, A
51	21	63.3	559	12	US-10-301-480-474416	Sequence 474416, A
52	21	63.3	559	12	US-10-301-480-1087825	Sequence 1087825, A
53	21	63.3	619	4	US-09-925-065A-42235	Sequence 42235, A
54	21	63.3	619	5	US-09-925-065A-42235	Sequence 42235, A
55	21	63.3	619	12	US-10-301-480-143473	Sequence 143473, A
56	21	63.3	619	12	US-10-301-480-756882	Sequence 756882, A
57	21	63.3	686	6	US-10-027-632-253967	Sequence 253967, A
58	21	63.3	686	7	US-10-027-632-253967	Sequence 253967, A
59	21	63.3	686	12	US-10-301-480-88975	Sequence 88975, A
60	21	63.3	686	12	US-10-301-480-702166	Sequence 702166, A
61	21	63.3	698	4	US-09-925-065A-63800	Sequence 63800, A
62	21	63.3	698	5	US-09-925-065A-63800	Sequence 63800, A
63	21	63.3	698	12	US-10-301-480-155038	Sequence 155038, A
64	21	63.3	698	12	US-10-301-480-778447	Sequence 778447, A
65	21	63.3	710	8	US-10-767-701-19896	Sequence 19896, A
66	21	63.3	1101	11	US-10-932-182A-82765	Sequence 82765, A
67	20.8	62.7	2256646	8	US-10-470-565-1	Sequence 1, Appl
68	20.4	61.4	1074	3	US-09-738-626-595	Sequence 595, App
69	20.4	61.4	1201	9	US-10-494-672-81	Sequence 81, Appl
70	20.4	61.4	2292	11	US-10-932-182A-78725	Sequence 78725, A
71	20.4	61.4	2780	8	US-10-451-467A-363	Sequence 363, App
72	20.4	61.4	3309400	3	US-09-738-626-1	Sequence 1, Appl
73	20.2	60.8	201	9	US-10-719-993-31292	Sequence 31292, A
74	20.2	60.8	1146	8	US-10-282-122A-25373	Sequence 25373, A
75	20.2	60.8	775062	10	US-10-719-993-6844	Sequence 6844, App
76	20.2	60.8	783062	9	US-10-461-862-166	Sequence 166, App
77	20	60.2	282	8	US-10-674-124A-11800	Sequence 11800, A
78	20	60.2	381	8	US-10-437-963-66738	Sequence 66738, A
79	20	60.2	560	4	US-09-925-065A-219816	Sequence 219816, A
80	20	60.2	560	4	US-09-925-065A-219817	Sequence 219817, A
81	20	60.2	560	4	US-09-925-065A-219818	Sequence 219818, A
82	20	60.2	560	5	US-09-925-065A-219816	Sequence 219816, A
83	20	60.2	560	5	US-09-925-065A-219817	Sequence 219817, A
84	20	60.2	560	5	US-09-925-065A-219818	Sequence 219818, A
85	20	60.2	575	12	US-10-301-480-305477	Sequence 305477, A
86	20	60.2	575	12	US-10-301-480-305478	Sequence 305478, A
87	20	60.2	575	12	US-10-301-480-305479	Sequence 305479, A
88	20	60.2	575	12	US-10-301-480-918886	Sequence 918886, A
89	20	60.2	575	12	US-10-301-480-918887	Sequence 918887, A
90	20	60.2	575	12	US-10-301-480-918888	Sequence 918888, A

C 91	20	60.2	669	7	US-10-312-273-340	Sequence 340, App	164	19.2	57.8	2952	16	US-11-136-527-4081	Sequence 4081, Ap
C 92	20	60.2	5161	3	US-09-764-869-1560	Sequence 1560, Ap	C 165	19.2	57.8	3096	8	US-10-282-122A-18427	Sequence 18427, A
C 93	20	60.2	5161	6	US-10-091-504-1560	Sequence 1560, Ap	166	19.2	57.8	3397	6	US-10-175-523-167	Sequence 167, App
C 94	20	60.2	5161	7	US-10-227-577-1560	Sequence 1560, Ap	167	19.2	57.8	3397	10	US-10-764-420-490	Sequence 490, App
C 95	20	60.2	1230025	7	US-10-289-762-1	Sequence 1, Appli	168	19.2	57.8	3397	13	US-11-099-266-167	Sequence 167, App
C 96	19.8	59.6	570	4	US-09-925-065A-519993	Sequence 519993, A	169	19.2	57.8	3477	8	US-10-152-319A-1414	Sequence 1414, Ap
C 97	19.8	59.6	570	5	US-09-925-065A-519993	Sequence 519993, A	170	19.2	57.8	3477	16	US-11-036-196-1414	Sequence 1414, Ap
C 98	19.8	59.6	611	6	US-10-027-632-212706	Sequence 212706, A	171	19.2	57.8	4090	13	US-11-097-143-16738	Sequence 16738, A
C 99	19.8	59.6	611	7	US-10-027-632-212706	Sequence 212706, A	172	19.2	57.8	12222	16	US-11-129-861-42	Sequence 42, Appli
C 100	19.8	59.6	1770	8	US-10-437-963-36811	Sequence 36811, A	173	19.2	57.8	25871	9	US-10-741-600-17660	Sequence 17660, A
C 101	19.6	59.0	356	10	US-10-950-009-184	Sequence 184, App	C 174	19.2	57.8	25871	10	US-10-995-561-13268	Sequence 13268, A
C 102	19.6	59.0	552	4	US-09-925-065A-159167	Sequence 159167, A	C 175	19.2	57.8	50679	9	US-10-719-993-6992	Sequence 6992, Ap
C 103	19.6	59.0	552	4	US-09-925-065A-159168	Sequence 159168, A	C 176	19.2	57.8	68193	11	US-10-330-773-377	Sequence 377, App
C 104	19.6	59.0	552	5	US-09-925-065A-159167	Sequence 159167, A	C 177	19.2	57.8	82660	6	US-10-037-192-2017	Sequence 2017, Ap
C 105	19.6	59.0	552	5	US-09-925-065A-159168	Sequence 159168, A	C 178	19.2	57.8	83493	9	US-10-331-053-61	Sequence 61, Appli
C 106	19.6	59.0	559	12	US-10-301-480-252093	Sequence 252093, A	C 179	19.2	57.8	83493	10	US-10-461-862-119	Sequence 119, App
C 107	19.6	59.0	559	12	US-10-301-480-252094	Sequence 252094, A	C 180	19.2	57.8	100000	15	US-11-124-368A-2898	Sequence 2898, Ap
C 108	19.6	59.0	559	12	US-10-301-480-865502	Sequence 865502, A	C 181	19.2	57.8	156321	12	US-10-960-414-483	Sequence 483, App
C 109	19.6	59.0	559	12	US-10-301-480-865503	Sequence 865503, A	C 182	19.2	57.8	183046	10	US-10-981-277-31	Sequence 31, Appli
C 110	19.6	59.0	602	4	US-09-925-065A-530489	Sequence 530489, A	C 183	19.2	57.8	215308	15	US-11-121-086-77	Sequence 77, Appli
C 111	19.6	59.0	602	4	US-09-925-065A-530490	Sequence 530490, A	C 184	19.2	57.8	322101	7	US-10-060-902-1	Sequence 1, Appli
C 112	19.6	59.0	602	4	US-09-925-065A-530491	Sequence 530491, A	C 185	19.2	57.8	322101	7	US-10-354-247-1	Sequence 1, Appli
C 113	19.6	59.0	602	5	US-09-925-065A-530489	Sequence 530489, A	C 186	19.2	57.8	1790242	9	US-10-719-993-6940	Sequence 6940, Ap
C 114	19.6	59.0	602	5	US-09-925-065A-530490	Sequence 530490, A	C 187	19.2	57.8	213	9	US-10-674-124A-841	Sequence 841, App
C 115	19.6	59.0	602	5	US-09-925-065A-530491	Sequence 530491, A	C 188	19.2	57.8	268	9	US-10-925-115-13896	Sequence 13896, A
C 116	19.6	59.0	622	4	US-09-925-065A-872774	Sequence 872774, A	C 189	19.2	57.8	460	4	US-09-925-065A-659347	Sequence 659347, A
C 117	19.6	59.0	622	5	US-09-925-065A-872774	Sequence 872774, A	C 190	19.2	57.8	460	5	US-09-925-065A-659347	Sequence 659347, A
C 118	19.6	59.0	660	8	US-10-767-701-7206	Sequence 7206, Ap	C 191	19.2	57.8	570	4	US-09-925-065A-490593	Sequence 490593, A
C 119	19.6	59.0	691	7	US-10-264-049-1868	Sequence 1868, Ap	C 192	19.2	57.8	570	4	US-09-925-065A-490595	Sequence 490595, A
C 120	19.6	59.0	765	8	US-10-424-599-141285	Sequence 141285, A	C 193	19.2	57.8	570	5	US-09-925-065A-490593	Sequence 490593, A
C 121	19.6	59.0	1083	8	US-10-425-114-33885	Sequence 33885, A	C 194	19.2	57.8	570	5	US-09-925-065A-490595	Sequence 490595, A
C 122	19.6	59.0	1305	8	US-10-425-114-19221	Sequence 19221, A	C 195	19.2	57.8	600	10	US-10-972-079-80034	Sequence 80034, A
C 123	19.6	59.0	1512	9	US-10-425-115-54675	Sequence 54675, A	C 196	19.2	57.8	656	6	US-10-027-632-270599	Sequence 270599, A
C 124	19.4	58.4	312	8	US-10-242-535A-2655	Sequence 2655, Ap	C 197	19.2	57.8	656	6	US-10-027-632-270600	Sequence 270600, A
C 125	19.4	58.4	312	8	US-10-085-783A-2655	Sequence 2655, Ap	C 198	19.2	57.8	656	7	US-10-027-632-270599	Sequence 270599, A
C 126	19.4	58.4	575	4	US-09-925-065A-533092	Sequence 533092, A	C 199	19.2	57.8	656	7	US-10-027-632-270600	Sequence 270600, A
C 127	19.4	58.4	575	4	US-09-925-065A-533093	Sequence 533093, A	C 200	19.2	57.8	729	8	US-10-767-701-20093	Sequence 20093, A
C 128	19.4	58.4	575	4	US-09-925-065A-533094	Sequence 533094, A	C 201	19.2	57.8	1918	8	US-10-424-599-69652	Sequence 69652, A
C 129	19.4	58.4	575	5	US-09-925-065A-533095	Sequence 533095, A	C 202	19.2	57.8	2232	8	US-10-282-122A-17006	Sequence 17006, A
C 130	19.4	58.4	575	5	US-09-925-065A-533092	Sequence 533092, A	C 203	19.2	57.8	2637	8	US-10-437-963-10512	Sequence 10512, A
C 131	19.4	58.4	575	5	US-09-925-065A-533093	Sequence 533093, A	C 204	19.2	57.8	3915	7	US-10-220-955-7	Sequence 7, Appli
C 132	19.4	58.4	575	5	US-09-925-065A-533094	Sequence 533094, A	C 205	19.2	57.8	4136	3	US-09-916-790-4	Sequence 4, Appli
C 133	19.4	58.4	575	5	US-09-925-065A-533095	Sequence 533095, A	C 206	19.2	57.8	4136	8	US-10-678-786-4	Sequence 4, Appli
C 134	19.4	58.4	991	12	US-10-301-480-561340	Sequence 561340, A	C 207	19.2	57.8	4151	9	US-10-357-930-24415	Sequence 24415, A
C 135	19.4	58.4	991	12	US-10-301-480-117479	Sequence 117479, A	C 208	19.2	57.8	4151	9	US-10-357-930-24455	Sequence 24455, A
C 136	19.4	58.4	992	10	US-10-795-159-11	Sequence 11, Appli	C 209	19.2	57.8	4359	8	US-10-437-963-10522	Sequence 10522, A
C 137	19.4	58.4	1757	8	US-10-296-115-506	Sequence 506, App	C 210	19.2	57.8	109565	8	US-10-322-281-27	Sequence 27, Appli
C 138	19.4	58.4	2416	8	US-10-302-172-392	Sequence 392, App	C 211	19.2	57.8	162173	15	US-11-121-086-72	Sequence 72, Appli
C 139	19.4	58.4	318760	9	US-10-719-993-6765	Sequence 6765, Ap	C 212	19.2	57.8	3186778	6	US-10-027-632-174961	Sequence 174961, A
C 140	19.2	57.8	50	16	US-11-175-859-84498	Sequence 84498, A	C 213	19.2	57.8	3186778	7	US-10-027-632-174961	Sequence 174961, A
C 141	19.2	57.8	201	9	US-10-741-600-36451	Sequence 36451, A	C 214	18.8	56.6	220	9	US-10-425-115-137592	Sequence 137592, A
C 142	19.2	57.8	201	10	US-10-995-561-34315	Sequence 34315, A	C 215	18.8	56.6	316	9	US-10-425-115-116492	Sequence 116492, A
C 143	19.2	57.8	418	3	US-09-792-246-1	Sequence 1, Appli	C 216	18.8	56.6	383	8	US-10-424-599-70502	Sequence 70502, A
C 144	19.2	57.8	418	3	US-10-425-115-176062	Sequence 176062, A	C 217	18.8	56.6	347	4	US-09-925-065A-551003	Sequence 551003, A
C 145	19.2	57.8	530	9	US-09-792-246-2	Sequence 2, Appli	C 218	18.8	56.6	547	5	US-09-925-065A-551003	Sequence 551003, A
C 146	19.2	57.8	583	3	US-09-792-246-2	Sequence 2, Appli	C 219	18.8	56.6	547	12	US-10-301-480-528953	Sequence 528953, A
C 147	19.2	57.8	583	8	US-10-424-599-77485	Sequence 77485, A	C 220	18.8	56.6	547	12	US-10-301-480-528953	Sequence 528953, A
C 148	19.2	57.8	869	7	US-10-437-963-20470	Sequence 20470, A	C 221	18.8	56.6	547	12	US-10-301-480-528953	Sequence 528953, A
C 149	19.2	57.8	1001	7	US-10-425-115-125727	Sequence 125727, A	C 222	18.8	56.6	579	5	US-09-925-065A-521724	Sequence 521724, A
C 150	19.2	57.8	1001	9	US-10-170-097-394	Sequence 394, App	C 223	18.8	56.6	579	5	US-09-925-065A-521724	Sequence 521724, A
C 151	19.2	57.8	1044	8	US-10-926-684-394	Sequence 394, App	C 224	18.8	56.6	594	10	US-10-779-543-17337	Sequence 17337, A
C 152	19.2	57.8	1334	4	US-10-437-963-99481	Sequence 99481, A	C 225	18.8	56.6	595	8	US-10-437-963-33530	Sequence 33530, A
C 153	19.2	57.8	1334	4	US-09-925-065A-718383	Sequence 718383, A	C 226	18.8	56.6	716	12	US-10-301-480-585657	Sequence 585657, A
C 154	19.2	57.8	1334	5	US-09-925-065A-718384	Sequence 718384, A	C 227	18.8	56.6	716	12	US-10-301-480-585657	Sequence 585657, A
C 155	19.2	57.8	1334	5	US-09-925-065A-718383	Sequence 718383, A	C 228	18.8	56.6	729	10	US-10-956-157-2836	Sequence 2836, Ap
C 156	19.2	57.8	1400	16	US-11-136-527-8177	Sequence 8177, Ap	C 229	18.8	56.6	729	10	US-10-956-157-2836	Sequence 2836, Ap
C 157	19.2	57.8	1828	3	US-09-841-260-6	Sequence 6, Appli	C 230	18.8	56.6	766	10	US-10-750-185-28863	Sequence 28863, A
C 158	19.2	57.8	1828	3	US-10-007-693-6	Sequence 6, Appli	C 231	18.8	56.6	766	10	US-10-750-185-28863	Sequence 28863, A
C 159	19.2	57.8	1828	6	US-10-007-693-6	Sequence 6, Appli	C 232	18.8	56.6	948	3	US-10-623-28863	Sequence 28863, A
C 160	19.2	57.8	1828	6	US-10-762-058-6	Sequence 6, Appli	C 233	18.8	56.6	948	3	US-10-623-28863	Sequence 28863, A
C 161	19.2	57.8	1828	10	US-10-197-220-6	Sequence 6, Appli	C 234	18.8	56.6	1017	11	US-10-932-182A-2567	Sequence 2567, A
C 162	19.2	57.8	1828	13	US-11-109-468-6	Sequence 6, Appli	C 235	18.8	56.6	1161	11	US-10-932-182A-2567	Sequence 2567, A
C 163	19.2	57.8	2090	13	US-11-097-143-16-1								

237	18.8	56.6	1400	13	US-11-060-756-124	Sequence 124, App	310	18.4	55.4	201	9	US-10-719-993-33929	Sequence 33929, A
238	18.8	56.6	1400	13	US-11-060-756-4396	Sequence 4396, App	311	18.4	55.4	201	9	US-10-741-600-72750	Sequence 72750, A
239	18.8	56.6	1611	16	US-11-096-568A-28699	Sequence 28699, A	312	18.4	55.4	201	9	US-10-741-600-72759	Sequence 72759, A
240	18.8	56.6	1620	16	US-11-096-568A-22492	Sequence 22492, A	313	18.4	55.4	320	9	US-10-767-795-5974	Sequence 5974, App
241	18.8	56.6	1749	8	US-10-437-963-8034	Sequence 8034, App	C 314	18.4	55.4	419	4	US-09-925-065A-555523	Sequence 555523, A
242	18.8	56.6	2055	10	US-10-750-185-46360	Sequence 46360, A	C 315	18.4	55.4	419	5	US-09-925-065A-555523	Sequence 555523, A
243	18.8	56.6	2055	10	US-10-750-623-46360	Sequence 46360, A	C 316	18.4	55.4	423	4	US-09-925-065A-273434	Sequence 273434, A
244	18.8	56.6	2871	11	US-10-932-182A-76779	Sequence 76779, A	C 317	18.4	55.4	423	5	US-09-925-065A-273434	Sequence 273434, A
245	18.8	56.6	3237	8	US-10-437-963-8037	Sequence 8037, App	C 318	18.4	55.4	425	3	US-09-814-353-15543	Sequence 15543, A
246	18.8	56.6	3540	8	US-10-437-963-7957	Sequence 7957, App	C 319	18.4	55.4	427	12	US-10-301-480-350658	Sequence 350658, A
247	18.8	56.6	4353	8	US-10-437-963-8028	Sequence 8028, App	C 320	18.4	55.4	427	12	US-10-301-480-350658	Sequence 350658, A
248	18.8	56.6	5143	8	US-10-437-963-8137	Sequence 8137, App	C 321	18.4	55.4	481	6	US-10-027-632-18592	Sequence 18592, A
249	18.8	56.6	5292	8	US-10-437-963-8086	Sequence 8086, App	C 322	18.4	55.4	481	7	US-10-027-632-18592	Sequence 18592, A
250	18.8	56.6	5337	8	US-10-437-963-8070	Sequence 8070, App	C 323	18.4	55.4	482	6	US-10-027-632-105785	Sequence 105785, A
251	18.8	56.6	5603	13	US-11-097-143-16049	Sequence 16049, A	C 324	18.4	55.4	482	7	US-10-027-632-105785	Sequence 105785, A
252	18.8	56.6	5976	8	US-10-437-963-8075	Sequence 8075, App	C 325	18.4	55.4	483	3	US-09-918-995-27013	Sequence 27013, A
253	18.8	56.6	7011	8	US-10-437-963-7966	Sequence 7966, App	C 326	18.4	55.4	483	9	US-10-767-795-5973	Sequence 5973, App
254	18.8	56.6	10327	13	US-11-097-143-16048	Sequence 16048, A	C 327	18.4	55.4	563	4	US-09-925-065A-642363	Sequence 642363, A
255	18.8	56.6	12000	8	US-10-311-507-102	Sequence 102, App	C 328	18.4	55.4	563	5	US-09-925-065A-642363	Sequence 642363, A
256	18.8	56.6	12963	10	US-10-486-319A-51	Sequence 51, Appl	329	18.4	55.4	579	4	US-09-925-065A-610506	Sequence 610506, A
257	18.8	56.6	12963	10	US-10-486-319A-73	Sequence 73, Appl	330	18.4	55.4	579	5	US-09-925-065A-610506	Sequence 610506, A
258	18.8	56.6	14362	8	US-10-433-793-78	Sequence 78, Appl	C 331	18.4	55.4	588	8	US-10-021-323-15155	Sequence 15155, A
259	18.8	56.6	14708	6	US-10-239-676-222	Sequence 222, App	C 332	18.4	55.4	590	4	US-09-925-065A-435575	Sequence 435575, A
260	18.8	56.6	14708	7	US-10-311-455-2218	Sequence 2218, App	C 333	18.4	55.4	590	5	US-09-925-065A-435575	Sequence 435575, A
261	18.8	56.6	14708	7	US-10-240-453-324	Sequence 324, App	C 334	18.4	55.4	597	12	US-10-301-480-497462	Sequence 497462, A
262	18.8	56.6	14708	8	US-10-221-714A-500	Sequence 500, App	C 335	18.4	55.4	597	12	US-10-301-480-1110871	Sequence 1110871, A
263	18.8	56.6	14708	8	US-10-240-589C-134	Sequence 134, App	C 336	18.4	55.4	598	9	US-10-767-795-5971	Sequence 5971, App
264	18.8	56.6	23483	10	US-10-461-862-43	Sequence 43, Appl	C 337	18.4	55.4	604	12	US-10-301-480-39026	Sequence 39026, A
265	18.8	56.6	106323	3	US-09-803-661-3	Sequence 3, Appl	C 338	18.4	55.4	604	12	US-10-301-480-652435	Sequence 652435, A
266	18.8	56.6	106323	6	US-10-300-827-3	Sequence 3, Appl	C 339	18.4	55.4	605	4	US-09-925-065A-610505	Sequence 610505, A
267	18.6	56.0	23	8	US-10-600-816-35	Sequence 35, Appl	C 340	18.4	55.4	605	5	US-09-925-065A-610505	Sequence 610505, A
268	18.6	56.0	400	2	US-08-781-986A-2945	Sequence 2945, App	C 341	18.4	55.4	610	4	US-09-925-065A-644481	Sequence 644481, A
269	18.6	56.0	400	8	US-10-329-624-2945	Sequence 2945, App	C 342	18.4	55.4	610	5	US-09-925-065A-644481	Sequence 644481, A
270	18.6	56.0	508	8	US-10-437-963-18619	Sequence 18619, A	C 343	18.4	55.4	613	12	US-10-301-480-364155	Sequence 364155, A
271	18.6	56.0	577	6	US-10-027-632-45110	Sequence 45110, A	C 344	18.4	55.4	616	4	US-10-301-480-977564	Sequence 977564, A
272	18.6	56.0	577	6	US-10-027-632-45110	Sequence 45110, A	C 345	18.4	55.4	616	5	US-09-925-065A-641333	Sequence 641333, A
273	18.6	56.0	580	6	US-10-027-632-310888	Sequence 310888, A	C 346	18.4	55.4	616	5	US-09-925-065A-641333	Sequence 641333, A
274	18.6	56.0	580	7	US-10-027-632-310888	Sequence 310888, A	C 347	18.4	55.4	618	4	US-09-925-065A-468645	Sequence 468645, A
275	18.6	56.0	582	6	US-10-027-632-273238	Sequence 273238, A	C 348	18.4	55.4	618	4	US-09-925-065A-468645	Sequence 468645, A
276	18.6	56.0	582	7	US-10-027-632-273238	Sequence 273238, A	C 349	18.4	55.4	618	4	US-09-925-065A-468645	Sequence 468645, A
277	18.6	56.0	593	9	US-09-925-065A-128629	Sequence 128629, A	C 350	18.4	55.4	618	5	US-09-925-065A-468645	Sequence 468645, A
278	18.6	56.0	596	4	US-09-925-065A-855691	Sequence 855691, A	C 351	18.4	55.4	618	5	US-09-925-065A-468645	Sequence 468645, A
279	18.6	56.0	596	5	US-09-925-065A-855691	Sequence 855691, A	C 352	18.4	55.4	618	5	US-09-925-065A-468645	Sequence 468645, A
280	18.6	56.0	598	4	US-09-925-065A-424942	Sequence 424942, A	C 353	18.4	55.4	624	6	US-10-027-632-224087	Sequence 224087, A
281	18.6	56.0	598	5	US-09-925-065A-424942	Sequence 424942, A	C 354	18.4	55.4	624	7	US-10-027-632-224087	Sequence 224087, A
282	18.6	56.0	600	10	US-10-972-079-8485	Sequence 8485, App	C 355	18.4	55.4	628	6	US-10-027-632-42035	Sequence 42035, A
283	18.6	56.0	600	10	US-10-972-079-8485	Sequence 8485, App	C 356	18.4	55.4	628	7	US-10-027-632-42035	Sequence 42035, A
284	18.6	56.0	608	12	US-10-301-480-488804	Sequence 488804, A	C 357	18.4	55.4	675	4	US-09-925-065A-844041	Sequence 844041, A
285	18.6	56.0	608	12	US-10-301-480-1102213	Sequence 1102213, A	C 358	18.4	55.4	675	5	US-09-925-065A-844041	Sequence 844041, A
286	18.6	56.0	906	10	US-10-499-353A-153	Sequence 153, App	C 359	18.4	55.4	676	4	US-09-925-065A-839712	Sequence 839712, A
287	18.6	56.0	934	8	US-10-425-114-31484	Sequence 31484, A	C 360	18.4	55.4	676	5	US-09-925-065A-839712	Sequence 839712, A
288	18.6	56.0	1011	10	US-10-750-185-40998	Sequence 40998, A	C 361	18.4	55.4	887	10	US-10-450-763-1953	Sequence 1953, App
289	18.6	56.0	1011	10	US-10-750-185-40998	Sequence 40998, A	C 362	18.4	55.4	887	10	US-10-450-763-1953	Sequence 1953, App
290	18.6	56.0	1094	8	US-10-424-599-70402	Sequence 70402, A	C 363	18.4	55.4	907	12	US-10-301-480-596334	Sequence 596334, A
291	18.6	56.0	1240	8	US-10-425-114-5783	Sequence 5783, App	C 364	18.4	55.4	910	8	US-10-360-708-9	Sequence 9, Appl
292	18.6	56.0	1284	8	US-10-282-122A-27069	Sequence 27069, App	C 365	18.4	55.4	924	12	US-10-301-480-535086	Sequence 535086, A
293	18.6	56.0	1351	10	US-10-764-420-22	Sequence 22, Appl	C 366	18.4	55.4	924	12	US-10-301-480-1148495	Sequence 1148495, A
294	18.6	56.0	1616	8	US-10-425-114-36385	Sequence 36385, A	C 367	18.4	55.4	1069	9	US-10-425-115-31464	Sequence 31464, A
295	18.6	56.0	1984	8	US-10-425-114-33768	Sequence 33768, A	C 368	18.4	55.4	1161	6	US-10-027-632-118067	Sequence 118067, A
296	18.6	56.0	2286	9	US-10-425-115-128635	Sequence 128635, A	C 369	18.4	55.4	1161	7	US-10-027-632-118067	Sequence 118067, A
297	18.6	56.0	2381	9	US-10-425-115-122443	Sequence 22443, A	C 370	18.4	55.4	1194	4	US-09-925-065A-287260	Sequence 287260, A
298	18.6	56.0	55606	7	US-10-085-117-133	Sequence 133, App	C 371	18.4	55.4	1194	5	US-09-925-065A-287260	Sequence 287260, A
299	18.6	56.0	67167	10	US-10-937-730A-3	Sequence 3, Appl	C 372	18.4	55.4	1276	10	US-10-750-185-48907	Sequence 48907, A
300	18.6	56.0	94750	8	US-10-672-787-38	Sequence 38, Appl	C 373	18.4	55.4	1276	10	US-10-750-623-48907	Sequence 48907, A
301	18.6	56.0	129908	6	US-10-270-875-1	Sequence 1, Appl	C 374	18.4	55.4	1297	4	US-09-925-065A-31739	Sequence 31739, A
302	18.6	56.0	129908	6	US-10-270-878-1	Sequence 1, Appl	C 375	18.4	55.4	1297	4	US-09-925-065A-31740	Sequence 31740, A
303	18.6	56.0	129908	6	US-10-270-786-1	Sequence 1, Appl	C 376	18.4	55.4	1297	5	US-09-925-065A-31739	Sequence 31739, A
304	18.6	56.0	129908	6	US-10-270-710-1	Sequence 1, Appl	C 377	18.4	55.4	1297	5	US-09-925-065A-31740	Sequence 31740, A
305	18.6	56.0	129908	6	US-10-270-859-1	Sequence 1, Appl	C 378	18.4	55.4	1297	6	US-10-027-632-254358	Sequence 254358, A
306	18.6	56.0	129908	7	US-10-270-846-1	Sequence 1, Appl	C 379	18.4	55.4	1297	7	US-10-027-632-254358	Sequence 254358, A
307	18.6	56.0	400660	9	US-10-388-838-68	Sequence 68, Appl	C 380	18.4	55.4	1297	12	US-10-301-480-89148	Sequence 89148, A
308	18.4	55.4	201	8	US-10-741-601-26008	Sequence 26008, A	C 381	18.4	55.4	1297	12	US-10-301-480-132977	Sequence 132977, A
309	18.4	55.4	201	8	US-10-741-601-26017	Sequence 26017, A	C 382	18.4	55.4	1297	12	US-10-301-480-132978	Sequence 132978, A

c 383	18.4	55.4	1297	12	US-10-301-480-702557	Sequence 702557,	456	18.2	54.8	4	US-09-925-065A-855054	Sequence 855054,
384	18.4	55.4	1297	12	US-10-301-480-746386	Sequence 746386,	457	18.2	54.8	5	US-09-925-065A-801239	Sequence 801239,
385	18.4	55.4	1297	12	US-10-301-480-746387	Sequence 746387,	458	18.2	54.8	5	US-09-925-065A-855054	Sequence 855054,
386	18.4	55.4	1561	16	US-11-096-566A-23073	Sequence 23073, A	c 459	18.2	54.8	541	US-10-301-480-490091	Sequence 490091,
387	18.4	55.4	1644	7	US-10-172-118-313	Sequence 313, App	c 460	18.2	54.8	541	US-10-301-480-1103500	Sequence 1103500,
388	18.4	55.4	1644	8	US-10-342-887-313	Sequence 313, App	c 461	18.2	54.8	544	US-09-925-065A-305958	Sequence 305958,
c 389	18.4	55.4	1812	4	US-09-925-065A-700542	Sequence 700542,	c 462	18.2	54.8	544	US-09-925-065A-305958	Sequence 305958,
c 390	18.4	55.4	1812	5	US-09-925-065A-700542	Sequence 700542,	c 463	18.2	54.8	548	US-10-301-480-381910	Sequence 381910,
c 391	18.4	55.4	2151	8	US-10-437-963-82702	Sequence 82702, A	c 464	18.2	54.8	548	US-10-301-480-995319	Sequence 995319,
c 392	18.4	55.4	2325	8	US-10-398-221-2279	Sequence 2279, App	c 465	18.2	54.8	551	US-09-925-065A-426637	Sequence 426637,
c 393	18.4	55.4	2376	8	US-10-398-221-672	Sequence 672, App	c 466	18.2	54.8	551	US-09-925-065A-426637	Sequence 426637,
c 394	18.4	55.4	2397	8	US-10-479-435-44	Sequence 44, Appl	c 467	18.2	54.8	552	US-09-925-065A-770573	Sequence 770573,
c 395	18.4	55.4	2485	8	US-10-425-114-5613	Sequence 5613, App	c 468	18.2	54.8	552	US-09-925-065A-770573	Sequence 770573,
c 396	18.4	55.4	2520	6	US-10-106-698-1327	Sequence 1327, App	c 469	18.2	54.8	552	US-10-437-963-61905	Sequence 61905, A
c 397	18.4	55.4	2526	9	US-10-425-115-13548	Sequence 113548,	c 470	18.2	54.8	556	US-09-925-065A-773105	Sequence 773105,
c 398	18.4	55.4	2830	6	US-10-098-841-196	Sequence 196, App	c 471	18.2	54.8	556	US-09-925-065A-773105	Sequence 773105,
c 399	18.4	55.4	2835	10	US-10-750-185-27523	Sequence 27523, A	c 472	18.2	54.8	560	US-09-925-065A-154707	Sequence 154707,
c 400	18.4	55.4	2835	10	US-10-750-623-27523	Sequence 27523, A	c 473	18.2	54.8	560	US-09-925-065A-154707	Sequence 154707,
c 401	18.4	55.4	3151	6	US-10-174-363-41	Sequence 41, Appl	c 474	18.2	54.8	560	US-09-925-065A-154709	Sequence 154709,
c 402	18.4	55.4	3151	13	US-11-093-888-41	Sequence 41, Appl	c 475	18.2	54.8	560	US-09-925-065A-154709	Sequence 154709,
c 403	18.4	55.4	4068	9	US-10-425-115-92048	Sequence 92048, A	c 476	18.2	54.8	560	US-09-925-065A-154708	Sequence 154708,
c 404	18.4	55.4	7930	13	US-11-097-143-13216	Sequence 13216, A	c 477	18.2	54.8	560	US-09-925-065A-154709	Sequence 154709,
c 405	18.4	55.4	14527	8	US-10-741-601-5617	Sequence 5617, App	c 478	18.2	54.8	566	US-09-925-065A-855465	Sequence 855465,
c 406	18.4	55.4	14527	9	US-10-741-600-17570	Sequence 17570, App	c 479	18.2	54.8	566	US-09-925-065A-855466	Sequence 855466,
c 407	18.4	55.4	18981	3	US-09-764-891-7649	Sequence 7649, App	c 480	18.2	54.8	566	US-09-925-065A-855465	Sequence 855465,
c 408	18.4	55.4	22463	8	US-10-741-601-5778	Sequence 5778, App	c 481	18.2	54.8	566	US-09-925-065A-855466	Sequence 855466,
c 409	18.4	55.4	22463	9	US-10-741-600-18006	Sequence 18006, A	c 482	18.2	54.8	572	US-10-301-480-248206	Sequence 248206,
c 410	18.4	55.4	31337	11	US-10-330-773-7	Sequence 7, Appli	c 483	18.2	54.8	572	US-10-301-480-248207	Sequence 248207,
c 411	18.4	55.4	43159	9	US-10-741-600-17953	Sequence 17953, A	c 484	18.2	54.8	572	US-10-301-480-248208	Sequence 248208,
c 412	18.4	55.4	81001	3	US-09-751-877-1	Sequence 1, Appli	c 485	18.2	54.8	572	US-10-301-480-861615	Sequence 861615,
c 413	18.4	55.4	81001	3	US-09-842-364-1	Sequence 1, Appli	c 486	18.2	54.8	572	US-10-301-480-861616	Sequence 861616,
c 414	18.4	55.4	81001	4	US-09-751-877-1	Sequence 1, Appli	c 487	18.2	54.8	572	US-10-301-480-861617	Sequence 861617,
c 415	18.4	55.4	81001	8	US-10-121-034-1	Sequence 1, Appli	c 488	18.2	54.8	592	US-09-925-065A-596588	Sequence 596588,
c 416	18.4	55.4	81001	9	US-10-121-034-1	Sequence 1, Appli	c 489	18.2	54.8	592	US-09-925-065A-596588	Sequence 596588,
c 417	18.4	55.4	161671	7	US-10-017-117-1	Sequence 1, Appli	c 490	18.2	54.8	600	US-10-972-079-82227	Sequence 82227, A
c 418	18.4	55.4	188971	8	US-10-235-192A-27	Sequence 27, Appl	c 491	18.2	54.8	600	US-10-972-079-85842	Sequence 85842, A
c 419	18.4	55.4	200000	8	US-10-672-764A-33	Sequence 33, Appl	c 492	18.2	54.8	600	US-10-972-079-85843	Sequence 85843, A
c 420	18.4	55.4	202251	6	US-10-087-192-985	Sequence 985, App	c 493	18.2	54.8	623	US-10-260-238-4982	Sequence 4982, Ap
c 421	18.4	55.4	277616	8	US-10-367-094-83	Sequence 83, Appl	c 494	18.2	54.8	633	US-09-925-065A-89255	Sequence 89255, A
c 422	18.4	55.4	370469	6	US-10-087-192-250	Sequence 250, App	c 495	18.2	54.8	633	US-09-925-065A-89256	Sequence 89256, A
c 423	18.4	55.4	495635	10	US-10-737-082-12	Sequence 12, Appl	c 496	18.2	54.8	633	US-09-925-065A-89257	Sequence 89257, A
c 424	18.4	55.4	495635	10	US-10-765-790-12	Sequence 12, Appl	c 497	18.2	54.8	633	US-09-925-065A-89255	Sequence 89255, A
c 425	18.4	55.4	546025	9	US-10-719-593-6862	Sequence 6862, App	c 498	18.2	54.8	633	US-09-925-065A-89256	Sequence 89256, A
c 426	18.4	55.4	705636	10	US-10-737-082-30	Sequence 30, Appl	c 499	18.2	54.8	633	US-09-925-065A-89257	Sequence 89257, A
c 427	18.4	55.4	705636	10	US-10-765-790-30	Sequence 30, Appl	c 500	18.2	54.8	633	US-10-301-480-190496	Sequence 190496,
c 428	18.4	55.4	1163020	8	US-10-398-221-10	Sequence 10, Appl	c 501	18.2	54.8	633	US-10-301-480-190497	Sequence 190497,
c 429	18.4	55.4	3011208	8	US-10-398-221-2058	Sequence 2058, App	c 502	18.2	54.8	633	US-10-301-480-190498	Sequence 190498,
c 430	18.2	54.8	157	8	US-10-424-599-40918	Sequence 40918, App	c 503	18.2	54.8	633	US-10-301-480-803905	Sequence 803905,
c 431	18.2	54.8	171	3	US-09-810-936-60	Sequence 60, Appl	c 504	18.2	54.8	633	US-10-301-480-803906	Sequence 803906,
c 432	18.2	54.8	171	3	US-09-429-755-60	Sequence 60, Appl	c 505	18.2	54.8	633	US-10-301-480-803907	Sequence 803907,
c 433	18.2	54.8	171	3	US-09-924-400-60	Sequence 60, Appl	c 506	18.2	54.8	635	US-10-027-632-286106	Sequence 286106,
c 434	18.2	54.8	171	7	US-10-212-679-60	Sequence 60, Appl	c 507	18.2	54.8	635	US-10-027-632-286106	Sequence 286106,
c 435	18.2	54.8	171	8	US-10-079-137B-60	Sequence 60, Appl	c 508	18.2	54.8	636	US-09-925-065A-504104	Sequence 504104,
c 436	18.2	54.8	171	16	US-11-139-041-60	Sequence 60, Appl	c 509	18.2	54.8	636	US-09-925-065A-504104	Sequence 504104,
c 437	18.2	54.8	176	3	US-09-864-761-22768	Sequence 22768, A	c 510	18.2	54.8	641	US-10-767-701-24640	Sequence 24640, A
c 438	18.2	54.8	201	9	US-10-719-993-12101	Sequence 12101, A	c 511	18.2	54.8	642	US-10-301-480-248209	Sequence 248209,
c 439	18.2	54.8	226	4	US-09-925-065A-542051	Sequence 542051, A	c 512	18.2	54.8	642	US-10-301-480-861618	Sequence 861618,
c 440	18.2	54.8	296	5	US-09-925-065A-542051	Sequence 542051, A	c 513	18.2	54.8	643	US-09-925-065A-154710	Sequence 154710,
c 441	18.2	54.8	356	9	US-10-425-115-165418	Sequence 165418, A	c 514	18.2	54.8	643	US-09-925-065A-154710	Sequence 154710,
c 442	18.2	54.8	356	3	US-09-918-995-37042	Sequence 37042, A	c 515	18.2	54.8	779	US-10-425-115-88731	Sequence 88731, A
c 443	18.2	54.8	400	8	US-10-442-535A-18317	Sequence 18317, A	c 516	18.2	54.8	825	US-10-027-632-161455	Sequence 161455,
c 444	18.2	54.8	400	8	US-10-085-783A-18317	Sequence 18317, A	c 517	18.2	54.8	825	US-10-027-632-161455	Sequence 161455,
c 445	18.2	54.8	449	6	US-10-779-543-20188	Sequence 20188, A	c 518	18.2	54.8	954	US-10-027-632-119708	Sequence 119708,
c 446	18.2	54.8	449	6	US-10-198-846-4859	Sequence 4859, App	c 519	18.2	54.8	954	US-10-027-632-119708	Sequence 119708,
c 447	18.2	54.8	470	3	US-09-864-761-6009	Sequence 6009, App	c 520	18.2	54.8	954	US-10-027-632-119707	Sequence 119707,
c 448	18.2	54.8	483	12	US-10-301-480-434183	Sequence 434183	c 521	18.2	54.8	954	US-10-027-632-119708	Sequence 119708,
c 449	18.2	54.8	483	12	US-10-301-480-1047592	Sequence 1047592,	c 522	18.2	54.8	1000	US-10-301-480-557602	Sequence 557602,
c 450	18.2	54.8	488	4	US-09-925-065A-363445	Sequence 363445,	c 523	18.2	54.8	1000	US-10-301-480-171011	Sequence 171011,
c 451	18.2	54.8	488	5	US-09-925-065A-363445	Sequence 363445,	c 524	18.2	54.8	1103	US-09-925-065A-724300	Sequence 724300,
c 452	18.2	54.8	502	3	US-09-783-590-4128	Sequence 4128, App	c 525	18.2	54.8	1103	US-09-925-065A-724300	Sequence 724300,
c 453	18.2	54.8	523	4	US-09-925-065A-810163	Sequence 810163,	c 526	18.2	54.8	1284	US-10-282-122A-20914	Sequence 20914, A
c 454	18.2	54.8	523	5	US-09-925-065A-810163	Sequence 810163,	c 527	18.2	54.8	1305	US-10-424-599-60455	Sequence 60455, A
c 455	18.2	54.8	526	4	US-09-925-065A-801239	Sequence 801239,	c 528	18.2	54.8	1328	US-10-425-115-87141	Sequence 87141, A

c 529	18.2	54.8	1594	6	US-10-084-817-129	Sequence 129, App	c 602	18	54.2	573	6	US-10-027-632-79811	Sequence 79811, A
c 530	18.2	54.8	1610	8	US-10-767-701-14213	Sequence 14213, A	c 603	18	54.2	573	6	US-10-027-632-80746	Sequence 80746, A
c 531	18.2	54.8	1622	8	US-10-424-599-114508	Sequence 114508, A	c 604	18	54.2	573	6	US-10-027-632-301330	Sequence 301330, A
c 532	18.2	54.8	2115	7	US-10-369-493-251077	Sequence 251077, A	c 605	18	54.2	573	7	US-10-027-632-79811	Sequence 79811, A
c 533	18.2	54.8	2118	11	US-10-932-182A-73440	Sequence 73440, A	c 606	18	54.2	573	7	US-10-027-632-80746	Sequence 80746, A
c 534	18.2	54.8	4365	7	US-10-240-965-46	Sequence 46, Appl	c 607	18	54.2	573	7	US-10-027-632-301330	Sequence 301330, A
c 535	18.2	54.8	29729	3	US-09-070-927A-238	Sequence 238, App	c 608	18	54.2	581	8	US-10-437-963-51798	Sequence 51798, A
c 536	18.2	54.8	64829	10	US-10-737-082-91	Sequence 91, Appl	c 609	18	54.2	589	4	US-09-925-065A-519202	Sequence 519202, A
c 537	18.2	54.8	64829	10	US-10-765-790-91	Sequence 91, Appl	c 610	18	54.2	589	4	US-09-925-065A-519203	Sequence 519203, A
c 538	18.2	54.8	127222	11	US-10-330-773-278	Sequence 278, App	c 611	18	54.2	589	5	US-09-925-065A-519203	Sequence 519203, A
c 539	18.2	54.8	380963	10	US-10-737-082-5	Sequence 5, Appl	c 612	18	54.2	589	5	US-09-925-065A-519203	Sequence 519203, A
c 540	18.2	54.8	380963	10	US-10-765-790-5	Sequence 5, Appl	c 613	18	54.2	600	10	US-10-972-079-6763	Sequence 6763, App
c 541	18.2	54.8	518360	8	US-10-367-094-125	Sequence 125, App	c 614	18	54.2	600	10	US-10-972-079-89184	Sequence 89184, A
c 542	18.2	54.8	606398	9	US-10-719-993-6782	Sequence 6782, App	c 615	18	54.2	600	10	US-10-972-079-89185	Sequence 89185, A
c 543	18.2	54.8	2731748	8	US-10-297-465A-1	Sequence 1, Appl	c 616	18	54.2	600	13	US-11-060-756-3557	Sequence 3557, App
c 544	18.2	54.8	9025608	7	US-10-156-761-1	Sequence 1, Appl	c 617	18	54.2	600	13	US-11-060-756-3557	Sequence 3557, App
c 545	18	54.2	188	9	US-10-425-115-152570	Sequence 152570, A	c 618	18	54.2	600	13	US-11-060-756-7828	Sequence 7828, App
c 546	18	54.2	230	6	US-10-066-543-1621	Sequence 1621, App	c 619	18	54.2	600	13	US-11-060-756-7829	Sequence 7829, App
c 547	18	54.2	242	15	US-11-148-108-16	Sequence 16, Appl	c 620	18	54.2	608	4	US-09-925-065A-2719	Sequence 2719, App
c 548	18	54.2	270	10	US-10-450-763-11091	Sequence 11091, A	c 621	18	54.2	608	5	US-09-925-065A-2719	Sequence 2719, App
c 549	18	54.2	270	10	US-10-450-763-11377	Sequence 11377, A	c 622	18	54.2	608	12	US-10-301-480-103956	Sequence 103956, A
c 550	18	54.2	270	10	US-10-450-763-12064	Sequence 12064, A	c 623	18	54.2	608	12	US-10-301-480-103956	Sequence 103956, A
c 551	18	54.2	298	9	US-10-674-124A-16915	Sequence 16915, A	c 624	18	54.2	668	4	US-09-925-065A-956854	Sequence 956854, A
c 552	18	54.2	320	9	US-10-357-930-47844	Sequence 47844, A	c 625	18	54.2	668	5	US-09-925-065A-956854	Sequence 956854, A
c 553	18	54.2	376	6	US-10-066-543-22	Sequence 22, Appl	c 626	18	54.2	753	6	US-10-027-632-150944	Sequence 150944, A
c 554	18	54.2	411	6	US-10-060-036-29975	Sequence 29975, App	c 627	18	54.2	753	7	US-10-027-632-150944	Sequence 150944, A
c 555	18	54.2	412	3	US-09-969-034-2789	Sequence 2789, App	c 628	18	54.2	812	6	US-10-027-632-171947	Sequence 171947, A
c 556	18	54.2	446	9	US-10-696-639-2906	Sequence 2906, App	c 629	18	54.2	812	7	US-10-027-632-171947	Sequence 171947, A
c 557	18	54.2	453	3	US-09-918-995-1492	Sequence 1492, App	c 630	18	54.2	831	10	US-10-750-185-60554	Sequence 60554, A
c 558	18	54.2	459	4	US-09-925-065A-202939	Sequence 202939, A	c 631	18	54.2	831	10	US-10-750-185-60554	Sequence 60554, A
c 559	18	54.2	459	4	US-09-925-065A-202940	Sequence 202940, A	c 632	18	54.2	842	8	US-09-988-292-8	Sequence 2779, App
c 560	18	54.2	459	5	US-09-925-065A-202939	Sequence 202939, A	c 633	18	54.2	878	3	US-09-988-292-8	Sequence 2779, App
c 561	18	54.2	459	5	US-09-925-065A-202940	Sequence 202940, A	c 634	18	54.2	878	8	US-10-776-601-8	Sequence 8, Appl
c 562	18	54.2	483	9	US-10-767-795-1780	Sequence 1780, App	c 635	18	54.2	906	8	US-10-776-601-8	Sequence 8, Appl
c 563	18	54.2	483	12	US-10-301-480-7709	Sequence 7709, App	c 636	18	54.2	906	9	US-10-425-114-25895	Sequence 11489, A
c 564	18	54.2	485	12	US-10-301-480-621118	Sequence 621118, A	c 637	18	54.2	907	9	US-10-425-115-62933	Sequence 62933, A
c 565	18	54.2	466	12	US-10-301-480-22645	Sequence 22645, A	c 638	18	54.2	992	12	US-10-301-480-591781	Sequence 591781, A
c 566	18	54.2	466	12	US-10-301-480-22646	Sequence 22646, A	c 639	18	54.2	992	12	US-10-301-480-591782	Sequence 591782, A
c 567	18	54.2	466	12	US-10-301-480-291283	Sequence 291283, A	c 640	18	54.2	992	12	US-10-301-480-1205190	Sequence 1205190, A
c 568	18	54.2	466	12	US-10-301-480-291284	Sequence 291284, A	c 641	18	54.2	992	12	US-10-301-480-1205191	Sequence 1205191, A
c 569	18	54.2	466	12	US-10-301-480-636054	Sequence 636054, A	c 642	18	54.2	996	12	US-10-301-480-577175	Sequence 577175, A
c 570	18	54.2	466	12	US-10-301-480-636055	Sequence 636055, A	c 643	18	54.2	996	12	US-10-301-480-577175	Sequence 577175, A
c 571	18	54.2	466	12	US-10-301-480-904692	Sequence 904692, A	c 644	18	54.2	1080	8	US-10-282-122A-7562	Sequence 7562, App
c 572	18	54.2	466	12	US-10-301-480-904693	Sequence 904693, A	c 645	18	54.2	1080	8	US-10-282-122A-7562	Sequence 7562, App
c 573	18	54.2	467	12	US-10-301-480-22647	Sequence 22647, A	c 646	18	54.2	1081	8	US-10-282-122A-35502	Sequence 35502, A
c 574	18	54.2	467	12	US-10-301-480-636056	Sequence 636056, A	c 647	18	54.2	1157	8	US-10-437-963-41074	Sequence 41074, A
c 575	18	54.2	476	8	US-10-424-599-15229	Sequence 15229, A	c 648	18	54.2	1165	10	US-10-450-763-11954	Sequence 11954, A
c 576	18	54.2	481	6	US-10-066-543-2792	Sequence 2792, App	c 649	18	54.2	1165	10	US-10-450-763-11954	Sequence 11954, A
c 577	18	54.2	481	6	US-10-060-036-2601	Sequence 2601, App	c 650	18	54.2	1197	8	US-10-437-963-51799	Sequence 51799, A
c 578	18	54.2	482	6	US-10-066-543-181	Sequence 181, App	c 651	18	54.2	1225	10	US-10-750-185-62836	Sequence 62836, A
c 579	18	54.2	482	6	US-10-066-543-1737	Sequence 1737, App	c 652	18	54.2	1225	10	US-10-750-185-62836	Sequence 62836, A
c 580	18	54.2	482	6	US-10-066-543-1898	Sequence 1898, App	c 653	18	54.2	1255	10	US-10-750-185-24712	Sequence 24712, A
c 581	18	54.2	483	6	US-10-066-543-2241	Sequence 2241, App	c 654	18	54.2	1255	10	US-10-750-623-24712	Sequence 24712, A
c 582	18	54.2	483	6	US-10-066-543-2794	Sequence 2794, App	c 655	18	54.2	1288	6	US-10-027-632-122922	Sequence 122922, A
c 583	18	54.2	491	6	US-10-027-632-96797	Sequence 96797, A	c 656	18	54.2	1288	7	US-10-027-632-122922	Sequence 122922, A
c 584	18	54.2	491	6	US-10-027-632-96798	Sequence 96798, A	c 657	18	54.2	1305	9	US-10-767-795-803	Sequence 803, App
c 585	18	54.2	491	6	US-10-027-632-96799	Sequence 96799, A	c 658	18	54.2	1341	8	US-10-282-122A-18356	Sequence 18356, App
c 586	18	54.2	491	7	US-10-027-632-96797	Sequence 96797, A	c 659	18	54.2	1400	13	US-11-060-756-2150	Sequence 2150, App
c 587	18	54.2	491	7	US-10-027-632-96798	Sequence 96798, A	c 660	18	54.2	1400	13	US-11-060-756-6422	Sequence 6422, App
c 588	18	54.2	491	7	US-10-027-632-96799	Sequence 96799, A	c 661	18	54.2	1407	7	US-10-437-963-101590	Sequence 101590, A
c 589	18	54.2	513	12	US-10-301-480-7710	Sequence 7710, App	c 662	18	54.2	1467	7	US-10-369-493-46250	Sequence 46250, A
c 590	18	54.2	513	12	US-10-301-480-621119	Sequence 621119, A	c 663	18	54.2	1512	7	US-10-305-720-850	Sequence 850, App
c 591	18	54.2	520	9	US-10-852-943-76	Sequence 76, Appl	c 664	18	54.2	1532	10	US-10-450-763-3946	Sequence 3946, App
c 592	18	54.2	524	3	US-09-998-598-2534	Sequence 2534, App	c 665	18	54.2	1532	6	US-10-044-205A-3	Sequence 3, Appl
c 593	18	54.2	541	9	US-10-653-047-1061	Sequence 1061, App	c 666	18	54.2	1662	3	US-09-964-469-1	Sequence 1, Appl
c 594	18	54.2	561	12	US-10-301-480-216342	Sequence 216342, A	c 667	18	54.2	1662	3	US-09-802-117-1	Sequence 1, Appl
c 595	18	54.2	561	12	US-10-301-480-216343	Sequence 216343, A	c 668	18	54.2	1662	3	US-10-217-745-1	Sequence 1, Appl
c 596	18	54.2	561	12	US-10-301-480-829751	Sequence 829751, A	c 669	18	54.2	1662	7	US-10-425-962-1	Sequence 1, Appl
c 597	18	54.2	561	12	US-10-301-480-829752	Sequence 829752, A	c 670	18	54.2	1662	8	US-10-311-034-45	Sequence 45, Appl
c 598	18	54.2	568	4	US-09-925-065A-117621	Sequence 117621, A	c 671	18	54.2	1662	9	US-10-451-168-49	Sequence 49, Appl
c 599	18	54.2	568	4	US-09-925-065A-117622	Sequence 117622, A	c 672	18	54.2	1662	9	US-10-788-197-22	Sequence 22, Appl
c 600	18	54.2	568	5	US-09-925-065A-117621	Sequence 117621, A	c 673	18	54.2	1662	10	US-10-980-387-49	Sequence 49, Appl
c 601	18	54.2	568	5	US-09-925-065A-117622	Sequence 117622, A	c 674	18	54.2	1662	11	US-10-984-548-1	Sequence 1, Appl

675	18	54.2	1662	16	US-11-040-218-22	Sequence 22, Appl	C 748	18	54.2	128993	9	US-10-484-577-681	Sequence 681, App
676	18	54.2	1683	15	US-11-148-108-19	Sequence 19, Appl	749	18	54.2	137560	9	US-10-481-112-1	Sequence 1, Appl
677	18	54.2	1701	8	US-10-072-012-273	Sequence 273, App	C 750	18	54.2	153752	11	US-10-330-773-508	Sequence 508, App
678	18	54.2	1705	8	US-10-424-599-13545	Sequence 13545, A	751	18	54.2	160921	6	US-10-087-192-1672	Sequence 1672, App
679	18	54.2	1888	8	US-10-437-963-64869	Sequence 64869, A	752	18	54.2	173308	10	US-10-756-149-629	Sequence 629, App
680	18	54.2	1950	9	US-10-472-928-3385	Sequence 3385, App	C 753	18	54.2	192392	10	US-10-461-862-95	Sequence 95, Appl
681	18	54.2	1953	8	US-10-474-776-5345	Sequence 545, App	C 754	18	54.2	314364	10	US-10-917-647-3	Sequence 3, Appl
682	18	54.2	1953	10	US-10-617-320-912	Sequence 912, App	C 755	18	54.2	387780	10	US-10-995-561-13259	Sequence 13259, A
683	18	54.2	2145	7	US-10-032-585-6532	Sequence 6532, App	C 756	18	54.2	519599	10	US-10-737-082-73	Sequence 73, Appl
684	18	54.2	2198	6	US-10-044-205A-1	Sequence 1, Appl	C 757	18	54.2	519599	10	US-10-765-790-73	Sequence 73, Appl
685	18	54.2	2249	3	US-09-802-117-5	Sequence 5, Appl	C 758	18	54.2	2162598	9	US-10-472-928-4579	Sequence 4979, App
686	18	54.2	2249	6	US-10-217-745-5	Sequence 5, Appl	C 759	17.8	53.6	160	3	US-09-908-975-20085	Sequence 20085, A
687	18	54.2	2249	11	US-10-984-548-5	Sequence 5, Appl	C 760	17.8	53.6	196	3	US-09-864-761-21524	Sequence 21524, A
688	18	54.2	2259	10	US-10-450-763-6124	Sequence 6124, App	C 761	17.8	53.6	242	7	US-10-026-925-52	Sequence 52, Appl
689	18	54.2	2371	3	US-09-754-862-4	Sequence 4, Appl	C 762	17.8	53.6	282	3	US-09-294-093B-203	Sequence 203, App
690	18	54.2	2371	6	US-10-256-590-4	Sequence 4, Appl	C 763	17.8	53.6	358	7	US-10-062-674-773	Sequence 773, App
691	18	54.2	2745	5	US-10-270-595-5	Sequence 5, Appl	C 764	17.8	53.6	367	3	US-09-864-761-4785	Sequence 4785, App
692	18	54.2	2745	9	US-10-772-437-5	Sequence 5, Appl	C 765	17.8	53.6	375	3	US-09-918-995-7710	Sequence 7710, App
693	18	54.2	2763	13	US-11-051-955-1	Sequence 1, Appl	C 766	17.8	53.6	407	3	US-09-960-352-13191	Sequence 13191, App
694	18	54.2	2813	15	US-11-148-108-18	Sequence 18, Appl	C 767	17.8	53.6	414	8	US-10-767-701-29417	Sequence 29417, A
695	18	54.2	2854	6	US-10-106-698-1971	Sequence 1971, App	C 768	17.8	53.6	426	8	US-10-424-599-133373	Sequence 133373, A
696	18	54.2	2854	11	US-10-081-088-11	Sequence 11, Appl	C 769	17.8	53.6	429	3	US-09-918-995-33930	Sequence 33930, A
697	18	54.2	2867	6	US-10-106-698-351	Sequence 351, App	C 770	17.8	53.6	431	7	US-10-062-674-395	Sequence 395, App
698	18	54.2	2880	8	US-10-437-963-83505	Sequence 83505, A	C 771	17.8	53.6	433	6	US-10-066-543-1837	Sequence 1837, App
699	18	54.2	2983	15	US-11-148-108-20	Sequence 20, Appl	C 772	17.8	53.6	438	6	US-10-027-632-88201	Sequence 88201, A
700	18	54.2	3007	6	US-10-055-412B-27	Sequence 27, Appl	C 773	17.8	53.6	438	7	US-10-027-632-88201	Sequence 88201, A
701	18	54.2	3007	9	US-10-779-949-27	Sequence 27, Appl	C 774	17.8	53.6	439	3	US-09-796-692-4516	Sequence 4516, App
702	18	54.2	3109	6	US-10-106-698-2111	Sequence 2111, App	C 775	17.8	53.6	439	6	US-10-040-862-4516	Sequence 4516, App
703	18	54.2	3111	3	US-09-823-356-25	Sequence 25, Appl	C 776	17.8	53.6	439	7	US-10-057-475B-4516	Sequence 4516, App
704	18	54.2	3111	3	US-09-981-353-191	Sequence 191, App	C 777	17.8	53.6	439	7	US-10-154-884B-4516	Sequence 4516, App
705	18	54.2	3111	6	US-10-235-994-25	Sequence 25, Appl	C 778	17.8	53.6	439	9	US-10-764-324-4516	Sequence 4516, App
706	18	54.2	3267	3	US-09-764-868-22	Sequence 22, Appl	C 779	17.8	53.6	459	3	US-09-770-444-464	Sequence 464, App
707	18	54.2	3311	3	US-09-922-217-1056	Sequence 1056, App	C 780	17.8	53.6	482	8	US-10-242-535A-37749	Sequence 37749, A
708	18	54.2	3311	6	US-09-833-263-1056	Sequence 1056, App	C 781	17.8	53.6	482	8	US-10-085-783A-37749	Sequence 37749, A
709	18	54.2	3311	6	US-10-025-380-1056	Sequence 1056, App	C 782	17.8	53.6	483	8	US-10-424-599-103211	Sequence 103211, A
710	18	54.2	3311	7	US-10-393-590-11	Sequence 11, Appl	C 783	17.8	53.6	499	4	US-09-925-065A-642257	Sequence 642257, App
711	18	54.2	3311	7	US-10-393-590-12	Sequence 12, Appl	C 784	17.8	53.6	499	5	US-09-925-065A-642256	Sequence 642256, App
712	18	54.2	3311	7	US-10-393-590-46	Sequence 46, Appl	C 785	17.8	53.6	516	4	US-09-925-065A-555038	Sequence 555038, App
713	18	54.2	3311	7	US-10-393-590-47	Sequence 47, Appl	C 786	17.8	53.6	516	5	US-09-925-065A-555038	Sequence 555038, App
714	18	54.2	3311	7	US-10-393-567-11	Sequence 11, Appl	C 787	17.8	53.6	531	10	US-10-958-216-229	Sequence 229, App
715	18	54.2	3311	7	US-10-393-567-12	Sequence 12, Appl	C 788	17.8	53.6	531	10	US-10-958-216-231	Sequence 231, App
716	18	54.2	3311	7	US-10-393-567-46	Sequence 46, Appl	C 789	17.8	53.6	560	4	US-09-925-065A-642257	Sequence 642257, App
717	18	54.2	3311	7	US-10-393-567-47	Sequence 47, Appl	C 790	17.8	53.6	560	5	US-09-925-065A-642257	Sequence 642257, App
718	18	54.2	3311	7	US-10-394-087-11	Sequence 11, Appl	C 791	17.8	53.6	574	12	US-10-301-480-611934	Sequence 611934, App
719	18	54.2	3311	7	US-10-394-087-12	Sequence 12, Appl	C 792	17.8	53.6	574	12	US-10-301-480-1225343	Sequence 1225343, App
720	18	54.2	3311	7	US-10-394-087-46	Sequence 46, Appl	C 793	17.8	53.6	575	4	US-09-925-065A-439823	Sequence 439823, App
721	18	54.2	3311	7	US-10-394-087-47	Sequence 47, Appl	C 794	17.8	53.6	575	5	US-09-925-065A-439823	Sequence 439823, App
722	18	54.2	3311	11	US-10-537-002-51	Sequence 51, Appl	C 795	17.8	53.6	576	4	US-09-925-065A-177481	Sequence 177481, App
723	18	54.2	3311	15	US-11-108-172-1056	Sequence 1056, App	C 796	17.8	53.6	576	4	US-09-925-065A-177482	Sequence 177482, App
724	18	54.2	3655	10	US-10-750-185-35242	Sequence 35242, A	C 797	17.8	53.6	576	4	US-09-925-065A-177483	Sequence 177483, App
725	18	54.2	3655	10	US-10-750-623-35242	Sequence 35242, A	C 798	17.8	53.6	576	4	US-09-925-065A-177484	Sequence 177484, App
726	18	54.2	3683	9	US-10-473-126-50	Sequence 50, Appl	C 799	17.8	53.6	576	5	US-09-925-065A-177481	Sequence 177481, App
727	18	54.2	3715	10	US-10-750-185-35377	Sequence 35377, A	C 800	17.8	53.6	576	5	US-09-925-065A-177482	Sequence 177482, App
728	18	54.2	3715	10	US-10-750-623-35377	Sequence 35377, A	C 801	17.8	53.6	576	5	US-09-925-065A-177483	Sequence 177483, App
729	18	54.2	4210	13	US-11-097-143-7852	Sequence 7852, App	C 802	17.8	53.6	576	5	US-09-925-065A-177484	Sequence 177484, App
730	18	54.2	4569	3	US-09-867-034-3	Sequence 3, Appl	C 803	17.8	53.6	578	12	US-10-301-480-500849	Sequence 500849, App
731	18	54.2	4569	3	US-10-276-115-3	Sequence 3, Appl	C 804	17.8	53.6	578	12	US-10-301-480-1114258	Sequence 1114258, App
732	18	54.2	4569	13	US-11-000-263-3	Sequence 3, Appl	C 805	17.8	53.6	581	12	US-10-301-480-521161	Sequence 521161, App
733	18	54.2	4712	2	US-08-961-527-120	Sequence 120, App	C 806	17.8	53.6	581	12	US-10-301-480-1134570	Sequence 1134570, App
734	18	54.2	7112	8	US-10-158-844-120	Sequence 120, App	C 807	17.8	53.6	585	4	US-09-925-065A-665548	Sequence 665548, App
735	18	54.2	7411	6	US-10-001-189-67	Sequence 67, Appl	C 808	17.8	53.6	585	5	US-09-925-065A-665548	Sequence 665548, App
736	18	54.2	10716	7	US-10-311-455-1391	Sequence 1391, App	C 809	17.8	53.6	588	4	US-09-925-065A-465058	Sequence 465058, App
737	18	54.2	13609	10	US-10-995-561-13442	Sequence 13442, A	C 810	17.8	53.6	588	5	US-09-925-065A-465058	Sequence 465058, App
738	18	54.2	17321	7	US-10-376-344-3	Sequence 3, Appl	C 811	17.8	53.6	592	4	US-09-925-065A-596587	Sequence 596587, App
739	18	54.2	32892	10	US-10-737-082-19	Sequence 19, Appl	C 812	17.8	53.6	592	4	US-09-925-065A-715181	Sequence 715181, App
740	18	54.2	32892	10	US-10-765-790-19	Sequence 19, Appl	C 813	17.8	53.6	592	5	US-09-925-065A-596587	Sequence 596587, App
741	18	54.2	34511	7	US-10-085-117-307	Sequence 307, App	C 814	17.8	53.6	592	5	US-09-925-065A-715181	Sequence 715181, App
742	18	54.2	36651	3	US-09-964-469-3	Sequence 3, Appl	C 815	17.8	53.6	597	4	US-09-925-065A-938280	Sequence 938280, App
743	18	54.2	36651	3	US-10-425-962-3	Sequence 3, Appl	C 816	17.8	53.6	597	5	US-09-925-065A-938280	Sequence 938280, App
744	18	54.2	44990	6	US-10-052-482-217	Sequence 217, App	C 817	17.8	53.6	600	4	US-09-925-065A-99481	Sequence 99481, A
745	18	54.2	61103	6	US-10-087-192-58	Sequence 58, Appl	C 818	17.8	53.6	600	5	US-09-925-065A-99481	Sequence 99481, A
746	18	54.2	78028	7	US-10-034-650-37	Sequence 37, Appl	C 819	17.8	53.6	600	10	US-10-972-075-67188	Sequence 67188, A
747	18	54.2	98472	9	US-10-484-577-673	Sequence 673, App	C 820	17.8	53.6	604	4	US-09-925-065A-694090	Sequence 694090, App

C 821	17.8	53.6	604	5	US-09-925-065A-694090	Sequence 694090,	C 894	17.8	53.6	1218	5	US-09-925-065A-36898	Sequence 36898, A
C 822	17.8	53.6	606	12	US-10-301-480-200183	Sequence 200183,	C 895	17.8	53.6	1218	5	US-09-925-065A-36899	Sequence 36899, A
C 823	17.8	53.6	606	12	US-10-301-480-813592	Sequence 813592,	C 896	17.8	53.6	1218	5	US-09-925-065A-36900	Sequence 36900, A
C 824	17.8	53.6	608	6	US-10-027-632-282271	Sequence 282271,	C 897	17.8	53.6	1218	5	US-09-925-065A-36901	Sequence 36901, A
C 825	17.8	53.6	608	7	US-10-027-632-282271	Sequence 282271,	C 898	17.8	53.6	1218	5	US-09-925-065A-36902	Sequence 36902, A
C 826	17.8	53.6	612	4	US-09-925-065A-412603	Sequence 412603,	C 899	17.8	53.6	1218	12	US-10-301-480-138135	Sequence 138135,
C 827	17.8	53.6	612	5	US-09-925-065A-412603	Sequence 412603,	C 900	17.8	53.6	1218	12	US-10-301-480-138136	Sequence 138136,
C 828	17.8	53.6	617	4	US-09-925-065A-449322	Sequence 449322,	C 901	17.8	53.6	1218	12	US-10-301-480-138137	Sequence 138137,
C 829	17.8	53.6	617	5	US-09-925-065A-449322	Sequence 449322,	C 902	17.8	53.6	1218	12	US-10-301-480-138138	Sequence 138138,
C 830	17.8	53.6	619	12	US-10-301-480-505466	Sequence 505466,	C 903	17.8	53.6	1218	12	US-10-301-480-138139	Sequence 138139,
C 831	17.8	53.6	619	12	US-10-301-480-604843	Sequence 604843,	C 904	17.8	53.6	1218	12	US-10-301-480-138140	Sequence 138140,
C 832	17.8	53.6	619	12	US-10-301-480-1118875	Sequence 1118875,	C 905	17.8	53.6	1218	12	US-10-301-480-751544	Sequence 751544,
C 833	17.8	53.6	619	12	US-10-301-480-1218252	Sequence 1218252,	C 906	17.8	53.6	1218	12	US-10-301-480-751545	Sequence 751545,
C 834	17.8	53.6	622	4	US-09-925-065A-445369	Sequence 445369,	C 907	17.8	53.6	1218	12	US-10-301-480-751546	Sequence 751546,
C 835	17.8	53.6	622	5	US-09-925-065A-445369	Sequence 445369,	C 908	17.8	53.6	1218	12	US-10-301-480-751547	Sequence 751547,
C 836	17.8	53.6	622	12	US-10-301-480-508808	Sequence 508808,	C 909	17.8	53.6	1218	12	US-10-301-480-751548	Sequence 751548,
C 837	17.8	53.6	629	12	US-10-301-480-1122217	Sequence 1122217,	C 910	17.8	53.6	1218	12	US-10-301-480-751549	Sequence 751549,
C 838	17.8	53.6	629	12	US-10-301-480-478680	Sequence 478680,	C 911	17.8	53.6	1218	12	US-09-981-876-98	Sequence 98, Appl
C 839	17.8	53.6	629	12	US-10-301-480-1092089	Sequence 1092089,	C 912	17.8	53.6	1253	3	US-09-148-545-98	Sequence 98, Appl
C 840	17.8	53.6	635	7	US-10-027-632-286105	Sequence 286105,	C 913	17.8	53.6	1253	10	US-10-979-111-98	Sequence 98, Appl
C 841	17.8	53.6	635	7	US-10-027-632-286105	Sequence 286105,	C 914	17.8	53.6	1352	13	US-11-097-143-21089	Sequence 21089, A
C 842	17.8	53.6	643	4	US-09-925-065A-743652	Sequence 743652,	C 915	17.8	53.6	1355	9	US-10-425-115-11229	Sequence 11229, A
C 843	17.8	53.6	643	5	US-09-925-065A-743652	Sequence 743652,	C 916	17.8	53.6	1404	6	US-10-027-632-258816	Sequence 258816,
C 844	17.8	53.6	645	3	US-09-989-534-1	Sequence 1, Appl	C 917	17.8	53.6	1404	7	US-10-027-632-258816	Sequence 258816,
C 845	17.8	53.6	652	4	US-09-925-065A-544663	Sequence 544663,	C 918	17.8	53.6	1404	12	US-10-301-480-93622	Sequence 93622, A
C 846	17.8	53.6	652	5	US-09-925-065A-544663	Sequence 544663,	C 919	17.8	53.6	1404	12	US-10-301-480-93622	Sequence 93622, A
C 847	17.8	53.6	659	9	US-10-425-115-47449	Sequence 47449, A	C 920	17.8	53.6	1409	10	US-10-750-185-34073	Sequence 34073, A
C 848	17.8	53.6	668	4	US-09-925-065A-746659	Sequence 746659,	C 921	17.8	53.6	1409	10	US-10-750-623-34073	Sequence 34073, A
C 849	17.8	53.6	668	5	US-09-925-065A-746659	Sequence 746659,	C 922	17.8	53.6	1428	3	US-09-925-301-467	Sequence 467, App
C 850	17.8	53.6	669	4	US-09-925-065A-805540	Sequence 805540,	C 923	17.8	53.6	1562	3	US-09-925-301-357	Sequence 357, App
C 851	17.8	53.6	669	5	US-09-925-065A-805540	Sequence 805540,	C 924	17.8	53.6	1564	10	US-10-501-282-285	Sequence 285, App
C 852	17.8	53.6	680	6	US-10-027-632-21912	Sequence 21912, A	C 925	17.8	53.6	1584	10	US-10-501-282-287	Sequence 287, App
C 853	17.8	53.6	680	6	US-10-027-632-21912	Sequence 21912, A	C 926	17.8	53.6	1584	10	US-10-501-282-289	Sequence 289, App
C 854	17.8	53.6	690	7	US-10-027-632-148774	Sequence 148774,	C 927	17.8	53.6	1584	10	US-10-501-282-291	Sequence 291, App
C 855	17.8	53.6	690	7	US-10-027-632-148774	Sequence 148774,	C 928	17.8	53.6	1599	8	US-10-437-963-57570	Sequence 57570, A
C 856	17.8	53.6	691	4	US-09-925-065A-807272	Sequence 807272,	C 929	17.8	53.6	1615	13	US-10-425-114-17802	Sequence 17802, A
C 857	17.8	53.6	691	5	US-09-925-065A-807272	Sequence 807272,	C 930	17.8	53.6	1615	13	US-11-097-143-39704	Sequence 39704, A
C 858	17.8	53.6	717	9	US-10-478-914-40	Sequence 40, Appl	C 931	17.8	53.6	1689	10	US-10-750-185-34070	Sequence 34070, A
C 859	17.8	53.6	725	10	US-10-750-185-50396	Sequence 50396, A	C 932	17.8	53.6	1689	10	US-10-750-623-34070	Sequence 34070, A
C 860	17.8	53.6	725	10	US-10-750-623-50396	Sequence 50396, A	C 933	17.8	53.6	1695	3	US-09-880-107-2276	Sequence 2276, Ap
C 861	17.8	53.6	773	3	US-09-981-876-34	Sequence 34, Appl	C 934	17.8	53.6	1695	10	US-10-989-891-25	Sequence 25, Appl
C 862	17.8	53.6	773	3	US-09-148-545-34	Sequence 34, Appl	C 935	17.8	53.6	1695	10	US-10-989-891-25	Sequence 99, Appl
C 863	17.8	53.6	773	10	US-10-979-111-34	Sequence 34, Appl	C 936	17.8	53.6	1995	8	US-10-424-599-38934	Sequence 38934, A
C 864	17.8	53.6	775	12	US-10-301-480-539608	Sequence 539608,	C 937	17.8	53.6	2002	8	US-10-424-599-38935	Sequence 38935, A
C 865	17.8	53.6	875	12	US-10-301-480-1153017	Sequence 1153017,	C 938	17.8	53.6	2085	9	US-10-425-115-165147	Sequence 165147, A
C 866	17.8	53.6	893	4	US-09-925-065A-94304	Sequence 94304, A	C 939	17.8	53.6	2115	16	US-10-282-122A-17703	Sequence 17703, A
C 867	17.8	53.6	893	5	US-09-925-065A-94304	Sequence 94304, A	C 940	17.8	53.6	2237	16	US-11-024-959-40	Sequence 40, Appl
C 868	17.8	53.6	893	12	US-10-301-480-195546	Sequence 195546,	C 941	17.8	53.6	2523	7	US-10-094-749-1277	Sequence 1277, Ap
C 869	17.8	53.6	893	12	US-10-301-480-808955	Sequence 808955,	C 942	17.8	53.6	2638	13	US-11-097-143-34948	Sequence 34948, A
C 870	17.8	53.6	911	8	US-10-425-114-30217	Sequence 30217, A	C 943	17.8	53.6	2648	7	US-10-104-047-349	Sequence 349, App
C 871	17.8	53.6	958	8	US-10-425-114-30214	Sequence 30214, A	C 944	17.8	53.6	2648	16	US-11-072-512-349	Sequence 349, App
C 872	17.8	53.6	959	8	US-10-425-114-30208	Sequence 30208, A	C 945	17.8	53.6	2676	11	US-10-932-182A-5102	Sequence 5102, Ap
C 873	17.8	53.6	960	12	US-10-301-480-578967	Sequence 578967,	C 946	17.8	53.6	2781	10	US-10-956-157-4775	Sequence 4775, Ap
C 874	17.8	53.6	960	12	US-10-301-480-1192376	Sequence 1192376,	C 947	17.8	53.6	2836	8	US-10-437-963-51222	Sequence 51222, A
C 875	17.8	53.6	998	12	US-10-301-480-551257	Sequence 551257,	C 948	17.8	53.6	2998	10	US-10-956-157-2547	Sequence 2547, Ap
C 876	17.8	53.6	998	12	US-10-301-480-1164666	Sequence 1164666,	C 949	17.8	53.6	3071	10	US-10-956-157-4776	Sequence 4776, Ap
C 877	17.8	53.6	999	12	US-10-301-480-582926	Sequence 582926,	C 950	17.8	53.6	3144	13	US-10-956-157-4776	Sequence 4776, Ap
C 878	17.8	53.6	1000	12	US-10-301-480-1196335	Sequence 1196335,	C 951	17.8	53.6	3234	7	US-11-097-143-23240	Sequence 23240, A
C 879	17.8	53.6	1000	12	US-10-301-480-538421	Sequence 538421,	C 952	17.8	53.6	3234	13	US-10-425-630-7	Sequence 7, Appl
C 880	17.8	53.6	1004	12	US-10-301-480-1151830	Sequence 1151830,	C 953	17.8	53.6	3234	13	US-11-133-849-7	Sequence 7, Appl
C 881	17.8	53.6	1022	9	US-10-437-963-82839	Sequence 82839, A	C 954	17.8	53.6	3483	8	US-10-437-963-55518	Sequence 55518, A
C 882	17.8	53.6	1022	9	US-10-425-115-49647	Sequence 49647, A	C 955	17.8	53.6	3483	8	US-10-437-963-55518	Sequence 55518, A
C 883	17.8	53.6	1040	6	US-10-198-846-7729	Sequence 7729, Ap	C 956	17.8	53.6	3724	13	US-11-097-143-21088	Sequence 21088, A
C 884	17.8	53.6	1137	8	US-10-425-114-13234	Sequence 13234, A	C 957	17.8	53.6	3724	13	US-11-097-143-21088	Sequence 21088, A
C 885	17.8	53.6	1182	3	US-09-764-891-8329	Sequence 8329, Ap	C 958	17.8	53.6	3935	16	US-11-097-143-7213	Sequence 7213, Ap
C 886	17.8	53.6	1183	3	US-09-764-891-8330	Sequence 8330, Ap	C 959	17.8	53.6	4024	16	US-11-203-033-1	Sequence 1, Appl
C 887	17.8	53.6	1218	4	US-09-925-065A-36897	Sequence 36897, A	C 960	17.8	53.6	4057	16	US-11-203-033-1	Sequence 1, Appl
C 888	17.8	53.6	1218	4	US-09-925-065A-36898	Sequence 36898, A	C 961	17.8	53.6	4083	9	US-10-357-930-23754	Sequence 23754, A
C 889	17.8	53.6	1218	4	US-09-925-065A-36899	Sequence 36899, A	C 962	17.8	53.6	4083	9	US-10-357-930-26557	Sequence 26557, A
C 890	17.8	53.6	1218	4	US-09-925-065A-36900	Sequence 36900, A	C 963	17.8	53.6	4360	13	US-11-097-143-23321	Sequence 23321, A
C 891	17.8	53.6	1218	4	US-09-925-065A-36901	Sequence 36901, A	C 964	17.8	53.6	5349	7	US-10-275-140-6	Sequence 6, Appl
C 892	17.8	53.6	1218	4	US-09-925-065A-36902	Sequence 36902, A	C 965	17.8	53.6	5412	7	US-10-275-140-6	Sequence 6, Appl
C 893	17.8	53.6	1218	5	US-09-925-065A-36897	Sequence 36897, A	C 966	17.8	53.6	5641	13	US-11-097-143-29239	Sequence 29239, A
											6	US-10-044-090-349	Sequence 349, App


```
967 17.8 53.6 5833 7 US-10-275-140-3
968 17.8 53.6 5965 6 US-10-215-050-1
969 17.8 53.6 6028 6 US-10-215-050-3
970 17.8 53.6 6028 7 US-10-275-140-1
971 17.8 53.6 6028 9 US-10-723-860-2269
972 17.8 53.6 6028 10 US-10-367-057-137
973 17.8 53.6 6099 9 US-10-723-860-6484
974 17.8 53.6 6446 16 US-11-136-527-543
975 17.8 53.6 7760 13 US-11-097-143-21082
976 17.8 53.6 7760 13 US-11-097-143-39703
977 17.8 53.6 7947 13 US-11-097-143-39703
978 17.8 53.6 8091 7 US-10-101-510-86
979 17.8 53.6 8091 7 US-10-356-625-1
980 17.8 53.6 8091 7 US-10-159-563-322
981 17.8 53.6 8091 8 US-10-301-832-4
982 17.8 53.6 8091 8 US-10-631-467-364
983 17.8 53.6 13295 13 US-11-097-143-11473
984 17.8 53.6 13341 2 US-08-910-386A-1
985 17.8 53.6 15743 7 US-10-240-453-269
986 17.8 53.6 18753 13 US-11-097-143-28264
987 17.8 53.6 21354 8 US-10-221-714A-512
988 17.8 53.6 27425 8 US-10-773-714-18
989 17.8 53.6 40568 6 US-10-087-192-1573
990 17.8 53.6 44348 8 US-10-301-832-11
991 17.8 53.6 52764 6 US-10-087-192-127
992 17.8 53.6 70123 6 US-10-087-192-610
993 17.8 53.6 96594 3 US-09-997-722-22
994 17.8 53.6 110079 13 US-11-099-266-96
995 17.8 53.6 160820 9 US-10-473-939-5
996 17.8 53.6 185750 16 US-11-114-798-56
997 17.8 53.6 196033 8 US-10-322-281-612
998 17.8 53.6 212231 6 US-10-087-192-1126
999 17.8 53.6 272022 11 US-10-330-773-102
c1000 17.8 53.6 304905 8 US-10-271-416-1
```

ALIGNMENTS

```
RESULT 1
US-10-600-816-33
; Sequence 33, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate Oligonucleotide.
; NAME/KEY: misc_feature
; LOCATION: (3)-(39)
; OTHER INFORMATION: wherein "n" equals A, C, G, or T.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)-(33)
; OTHER INFORMATION: wherein "y" equals C, or T.
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (27)...(42)
; OTHER INFORMATION: wherein "x" equals A, or G.
US-10-600-816-33
Query Match 100.0%; Score 33.2; DB 8; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
RESULT 2
US-10-600-816-30
; Sequence 30, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-30
Query Match 90.4%; Score 30; DB 8; Length 42;
Best Local Similarity 64.3%; Pred. No. 0.002;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
Db 1 GCCCAGCTTGGCCGAGCCCTTACAAGACTATGAAGTAAAG 42
RESULT 3
US-10-066-543-1937
; Sequence 1937, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Iodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1937
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-10-066-543-1937

Query Match 90.4%; Score 30; DB 6; Length 497;
Best Local Similarity 64.3%; Pred. No. 0.004;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 344 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 385

RESULT 4

US-10-066-543-1811/c
; Sequence 1811, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Barrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066.543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1811

Query Match 90.4%; Score 30; DB 6; Length 552;
Best Local Similarity 64.3%; Pred. No. 0.0041;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCNCNTAYAARGAYTAYGAGTNAAR 42
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 154 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 113

RESULT 5

US-09-969-034-2222/c
; Sequence 2222, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969.034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2222
; LENGTH: 620

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 491, 500, 517, 530, 545, 547, 553, 558, 570, 599, 620

; OTHER INFORMATION: n = A,T,C or G

US-09-969-034-2222

Query Match

90.4%; Score 30; DB 3; Length 620;

Best Local Similarity

64.3%; Pred. No. 0.0043;

Matches

27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCNCNTAYAARGAYTAYGAGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

Db 154 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 113

RESULT 6

US-09-969-034-3393/c
; Sequence 3393, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969.034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3393
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 363, 470, 472, 516, 523, 531, 551, 567, 573, 588, 594, 613,
; LOCATION: 623
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3393

Query Match

90.4%; Score 30; DB 3; Length 634;

Best Local Similarity

64.3%; Pred. No. 0.0043;

Matches

27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYGCGTGGCCNTCNCNTAYAARGAYTAYGAGTNAAR 42

|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

Db 158 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 117

RESULT 7

US-10-125-968-701/c
; Sequence 701, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER


```
; FILE REFERENCE: MRI-032
; CURRENT APPLICATION NUMBER: US/10/125,968
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,163
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 701
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 72, 76, 77, 99, 101, 423, 445, 468, 489, 557, 566, 633
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-701

Query Match          90.4%; Score 30; DB 7; Length 642;
Best Local Similarity 64.3%; Pred. No. 0.0043;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 143

RESULT 8
US-10-712-615-134
; Sequence 134, Application US/10712615
; Publication No. US20040214317A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY8, EXPRESSED
; FILE REFERENCE: D0047A-CIP
; CURRENT APPLICATION NUMBER: US/10/712,615
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: U.S. 09/992,238
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: U.S. 60/248,285
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: U.S. 60/268,581
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: U.S. 60/308,285
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: U.S. 60/317,166
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 134
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: dna;Homo sapiens
US-10-712-615-134

Query Match          90.4%; Score 30; DB 9; Length 1071;
Best Local Similarity 64.3%; Pred. No. 0.005;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1018 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1059

RESULT 9
US-09-866-050A-249
; Sequence 249, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
```

```
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-09-866-050A-249

Query Match          90.4%; Score 30; DB 3; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.0051;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 990 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 10
US-10-152-661-249
; Sequence 249, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Human
US-10-152-661-249

Query Match          90.4%; Score 30; DB 6; Length 1212;
Best Local Similarity 64.3%; Pred. No. 0.0051;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGNTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 990 GCCCAGCGTTGGCGAGCCCTTACAAAGACTATGAAGTAAAG 1031

RESULT 11
US-10-313-542-223
```

```
/ Sequence 223, Application US/10313542
/ Publication No. US20030120057A1
/ GENERAL INFORMATION:
/ APPLICANT: Roopa, Reddy
/ APPLICANT: Guegler, Karl, J.
/ APPLICANT: Au-Young, Janice
/ TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
/ FILE REFERENCE: PA-0013 US
/ CURRENT APPLICATION NUMBER: US/10/313,542
/ CURRENT FILING DATE: 2002-12-05
/ PRIOR APPLICATION NUMBER: US/09/495,050
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/118,318
/ PRIOR FILING DATE: 1999-02-01
/ NUMBER OF SEQ ID NOS: 305
/ SOFTWARE: PERL Program
/ SEQ ID NO 223
/ LENGTH: 1228
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030120057A1 2681738CT1
US-10-313-542-223

Query Match          90.4%; Score 30; DB 7; Length 1228;
Best Local Similarity 64.3%; Pred. No. 0.0052;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
   |||:||||| |||:||||| |||:||||| |||:|||||
Db 541 GCCACGCTTGGCCGAGCCCTTACAAGAGACTATGAAGTAAAG 582

RESULT 12
US-11-060-756-2418
/ Sequence 2418, Application US/11060756
/ Publication No. US20050221354A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William Martin
/ TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
/ FILE OF INVENTION: Target Genes
/ FILE REFERENCE: AM101083 (031896-042000)
/ CURRENT APPLICATION NUMBER: US/11/060,756
/ CURRENT FILING DATE: 2005-02-18
/ NUMBER OF SEQ ID NOS: 303284
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2418
/ LENGTH: 1400
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-060-756-2418

Query Match          90.4%; Score 30; DB 13; Length 1400;
Best Local Similarity 64.3%; Pred. No. 0.0054;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
   |||:||||| |||:||||| |||:||||| |||:|||||
Db 215 GCCACGCTTGGCCGAGCCCTTACAAGAGACTATGAAGTAAAG 256

RESULT 13
US-11-060-756-6690
/ Sequence 6690, Application US/11060756
/ Publication No. US20050221354A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William Martin
/ TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
/ FILE OF INVENTION: Target Genes
/ FILE REFERENCE: AM101083 (031896-042000)
```

```
/ CURRENT APPLICATION NUMBER: US/11/060,756
/ CURRENT FILING DATE: 2005-02-18
/ NUMBER OF SEQ ID NOS: 303284
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 6690
/ LENGTH: 1400
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-060-756-6690

Query Match          90.4%; Score 30; DB 13; Length 1400;
Best Local Similarity 64.3%; Pred. No. 0.0054;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
   |||:||||| |||:||||| |||:||||| |||:|||||
Db 215 GCCACGCTTGGCCGAGCCCTTACAAGAGACTATGAAGTAAAG 256

RESULT 14
US-10-936-626-64
/ Sequence 64, Application US/10936626
/ Publication No. US20050106644A1
/ GENERAL INFORMATION:
/ APPLICANT: Cairns, Belinda
/ APPLICANT: Chen, Ruihuan
/ APPLICANT: Frantz, Gretchen
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Koepfen, Hartmut
/ APPLICANT: Phillips, Heidi S.
/ APPLICANT: Polakis, Paul
/ APPLICANT: Spencer, Susan D.
/ APPLICANT: Smith, Victoria
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wu, Thomas D.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
/ FILE OF INVENTION: Treatment of Tumor
/ FILE REFERENCE: P5001R1P1
/ CURRENT APPLICATION NUMBER: US/10/936,626
/ CURRENT FILING DATE: 2004-09-08
/ PRIOR APPLICATION NUMBER: US 10/872,991
/ PRIOR FILING DATE: 2004-06-21
/ PRIOR APPLICATION NUMBER: US 10/872,972
/ PRIOR FILING DATE: 2004-06-21
/ PRIOR APPLICATION NUMBER: US 10/241,220
/ PRIOR FILING DATE: 2002-09-11
/ PRIOR APPLICATION NUMBER: US 10/177,488
/ PRIOR FILING DATE: 2002-06-19
/ PRIOR APPLICATION NUMBER: US 60/299,500
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/301,880
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: US 60/323,268
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: US 60/557,116
/ PRIOR FILING DATE: 2004-03-26
/ PRIOR APPLICATION NUMBER: US 60/598,899
/ PRIOR FILING DATE: 2004-08-04
/ NUMBER OF SEQ ID NOS: 154
/ SEQ ID NO 64
/ LENGTH: 1460
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-936-626-64

Query Match          90.4%; Score 30; DB 10; Length 1460;
Best Local Similarity 64.3%; Pred. No. 0.0054;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAARGAYTAYGARGTNAAR 42
   |||:||||| |||:||||| |||:||||| |||:|||||
Db 1145 GCCACGCTTGGCCGAGCCCTTACAAGAGACTATGAAGTAAAG 1186
```

```
RESULT 15
US-10-938-061-64
; Sequence 64, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairne, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; FILE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1F1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 64
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-938-061-64

Query Match          90.4%; Score 30; DB 10; Length 1460;
Best Local Similarity 64.3%; Pred. No. 0.0054;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||:||||| |||:||||| |||:||||| |||:|||||
Db 1145 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1186

RESULT 16
US-10-224-289-5
; Sequence 5, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID A.
; APPLICANT: STEWART, TIMOTHY A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; FILE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
```

```
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-289-5

Query Match          90.4%; Score 30; DB 7; Length 1619;
Best Local Similarity 64.3%; Pred. No. 0.0056;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||:||||| |||:||||| |||:||||| |||:|||||
Db 1137 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1178

RESULT 17
US-10-935-190-43
; Sequence 43, Application US/10935190
; Publication No. US20050037466A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Valda
; APPLICANT: BURFORD, Neil
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; TITLE OF INVENTION: RECEPTORS AND ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0726 PCT
; CURRENT APPLICATION NUMBER: US/10/935,190
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US/10/031,904
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/145,232; 60/158,578; 60/165,192
; PRIOR FILING DATE: 1999-07-21; 1999-10-07; 1999-11-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 43
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2681738CB1
US-10-935-190-43

Query Match          90.4%; Score 30; DB 9; Length 1619;
Best Local Similarity 64.3%; Pred. No. 0.0056;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||:||||| |||:||||| |||:||||| |||:|||||
Db 1137 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1178

RESULT 18
US-10-505-486-196
; Sequence 196, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
```

```
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 196
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Human
; US-10-505-486-196

Query Match          90.4%; Score 30; DB 10; Length 1788;
Best Local Similarity 64.3%; Pred. No. 0.0058;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 1018 GCCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAAG 1059

RESULT 19
US-10-775-920-13
; Sequence 13, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-13

Query Match          90.4%; Score 30; DB 9; Length 2297;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 1125 GCCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAAG 1166

RESULT 20
US-10-224-289-3
; Sequence 3, Application US/10224289
; Publication No. US20030207288A1
; GENERAL INFORMATION:
; APPLICANT: STEWART, TIMOTHY A.
; APPLICANT: LEWIN, DAVID A.
; TITLE OF INVENTION: GPCR-LIKE RETINOIC ACID-INDUCED GENE 1 PROTEIN AND
; TITLE OF INVENTION: NUCLEIC ACID
; FILE REFERENCE: 9800081-0085
; CURRENT APPLICATION NUMBER: US/10/224,289
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,940
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3
; LENGTH: 2302
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; US-10-224-289-3

Query Match          90.4%; Score 30; DB 7; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 1117 GCCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAAG 1158

RESULT 21
US-10-240-425-405
; Sequence 405, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 405
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF095448
; US-10-240-425-405

Query Match          90.4%; Score 30; DB 8; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 1117 GCCACGCTTGGCGGAGCCCTTACAAGACTATGAAGTAAG 1158

RESULT 22
US-10-775-920-9
; Sequence 9, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-9
```

```
Query Match          90.4%; Score 30; DB 9; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 1117 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 23
US-10-510-507-2
; Sequence 2, Application US/10510507
; Publication No. US20050282165A1
; GENERAL INFORMATION:
; APPLICANT: Terrett, Jonathan A
; TITLE OF INVENTION: DIAGNOSIS OF CARCINOMA USING RAIG1 POLYPEPTIDES
; FILE REFERENCE: 2543-1-039PCT/US
; CURRENT APPLICATION NUMBER: US/10/510,507
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: GB0208331.9
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: GB0221538.2
; PRIOR FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-510-507-2

Query Match          90.4%; Score 30; DB 10; Length 2302;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 1117 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1158

RESULT 24
US-10-775-920-12
; Sequence 12, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 12
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-12

Query Match          90.4%; Score 30; DB 9; Length 2305;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 1121 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1162

RESULT 25
US-10-176-847-59
```

```
; Sequence 59, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-59

Query Match          90.4%; Score 30; DB 6; Length 2316;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 1123 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1164

RESULT 26
US-11-080-991-59
; Sequence 59, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-080-991-59

Query Match          90.4%; Score 30; DB 15; Length 2316;
Best Local Similarity 64.3%; Pred. No. 0.0062;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGNTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
Db 1123 GCCACGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1164

RESULT 27
US-10-775-920-11
; Sequence 11, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 395
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 2446
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-11

Query Match      90.4%; Score 30; DB 9; Length 2446;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGAGCTNAAR 42
Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 28
US-10-225-567A-453
; Sequence 453, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 453
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-225-567A-453

Query Match      90.4%; Score 30; DB 6; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGAGCTNAAR 42
Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 29
US-10-269-909-63
; Sequence 63, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBIZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-269-909-63
```

```
Query Match      90.4%; Score 30; DB 7; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGAGCTNAAR 42
Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 30
US-10-269-909-64
; Sequence 64, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBIZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-269-909-64

Query Match      90.4%; Score 30; DB 7; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAAAGAYTAYGAGCTNAAR 42
Db 1271 GCCACGCTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 31
US-10-295-027-619
; Sequence 619, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
```

```
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-619

Query Match          90.4%; Score 30; DB 7; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db 1271 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 32
US-10-600-816-2
; Sequence 2, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-600-816-2

Query Match          90.4%; Score 30; DB 8; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db 1271 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 33
US-10-600-816-18
; Sequence 18, Application US/10600816
; Publication No. US20040121362A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION AND MODULATION OF A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: (GPCR), RAI3, ASSOCIATED WITH CHRONIC OBSTRUCTIVE PULMONARY
; TITLE OF INVENTION: DISEASE (COPD) AND NF-KB AND E-SELECTIN REGULATION
; FILE REFERENCE: D0251 NP
; CURRENT APPLICATION NUMBER: US/10/600,816
; CURRENT FILING DATE: 2003-06-20
```

```
; PRIOR APPLICATION NUMBER: U.S. 60/390,850
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: U.S. 60/407,006
; PRIOR FILING DATE: 2002-08-29
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RAI3 Polymorphic Allele Summary Sequence.
; NAME/KEY: misc feature
; LOCATION: (112)..(112)
; OTHER INFORMATION: wherein "n" equals either G or A.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (364)..(364)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: wherein "n" equals either C or T.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: wherein "n" equals either A or G.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1111)..(1111)
; OTHER INFORMATION: wherein "n" equals either T or C.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1173)..(1173)
; OTHER INFORMATION: wherein "n" equals either A or G.
US-10-600-816-18

Query Match          90.4%; Score 30; DB 8; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGARGTNAAR 42
    |||||
Db 1271 GCCCAGCGTTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 34
US-10-775-920-10
; Sequence 10, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 395
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2456
```


; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-10

Query Match 90.4%; Score 30; DB 9; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCGTGGCCNTCCNTAARGAYTAYGAGTNAAR 42
|||:||||| |||:|||||:||||:||||:
Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 35

US-10-936-626-40
; Sequence 40, Application US/10936626
; Publication No. US20050106644A1

; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda

; APPLICANT: Chen, Ruihuan

; APPLICANT: Frantz, Gretchen

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Koepfen, Hartmut

; APPLICANT: Phillips, Heidi S.

; APPLICANT: Polakis, Paul

; APPLICANT: Spencer, Susan D.

; APPLICANT: Smith, Victoria

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wu, Thomas D.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor

; FILE REFERENCE: P5001R1P1

; CURRENT APPLICATION NUMBER: US/10/936,626

; PRIOR FILING DATE: 2004-09-08

; PRIOR APPLICATION NUMBER: US 10/872,991

; PRIOR FILING DATE: 2004-06-21

; PRIOR APPLICATION NUMBER: US 10/872,972

; PRIOR FILING DATE: 2004-06-21

; PRIOR APPLICATION NUMBER: US 10/241,220

; PRIOR FILING DATE: 2002-09-11

; PRIOR APPLICATION NUMBER: US 10/177,488

; PRIOR FILING DATE: 2002-06-19

; PRIOR APPLICATION NUMBER: US 60/299,500

; PRIOR FILING DATE: 2001-06-20

; PRIOR APPLICATION NUMBER: US 60/301,880

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 60/323,268

; PRIOR FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US 60/557,116

; PRIOR FILING DATE: 2004-03-26

; PRIOR APPLICATION NUMBER: US 60/598,899

; PRIOR FILING DATE: 2004-08-04

; NUMBER OF SEQ ID NOS: 154

; SEQ ID NO 40

; LENGTH: 2456

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-936-626-40

Query Match 90.4%; Score 30; DB 10; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCGTGGCCNTCCNTAARGAYTAYGAGTNAAR 42
|||:||||| |||:|||||:||||:||||:
Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 36

US-10-938-061-40

; Sequence 40, Application US/10938061

; Publication No. US20050107595A1

; GENERAL INFORMATION:

; APPLICANT: Cairns, Belinda

; APPLICANT: Chen, Ruihuan

; APPLICANT: Frantz, Gretchen

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Koepfen, Hartmut

; APPLICANT: Phillips, Heidi S.

; APPLICANT: Polakis, Paul

; APPLICANT: Spencer, Susan D.

; APPLICANT: Smith, Victoria

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wu, Thomas D.

; APPLICANT: Zhang, Zemin

; APPLICANT: Sakanaka, Chie

; APPLICANT: Chuntharapai, Anan

; APPLICANT: Reed Chae J.

; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor

; FILE REFERENCE: P5001R1P1B

; CURRENT APPLICATION NUMBER: US/10/938,061

; PRIOR FILING DATE: 2004-09-10

; PRIOR APPLICATION NUMBER: US 10/872,991

; PRIOR FILING DATE: 2004-06-21

; PRIOR APPLICATION NUMBER: US 10/872,972

; PRIOR FILING DATE: 2004-06-21

; PRIOR APPLICATION NUMBER: US 10/241,220

; PRIOR FILING DATE: 2002-09-11

; PRIOR APPLICATION NUMBER: US 10/177,488

; PRIOR FILING DATE: 2002-06-19

; PRIOR APPLICATION NUMBER: US 60/299,500

; PRIOR FILING DATE: 2001-06-20

; PRIOR APPLICATION NUMBER: US 60/301,880

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 60/323,268

; PRIOR FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US 60/557,116

; PRIOR FILING DATE: 2004-03-26

; PRIOR APPLICATION NUMBER: US 60/598,899

; PRIOR FILING DATE: 2004-08-04

; NUMBER OF SEQ ID NOS: 154

; SEQ ID NO 40

; LENGTH: 2456

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-938-061-40

Query Match 90.4%; Score 30; DB 10; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GCNCAYCGTGGCCNTCCNTAARGAYTAYGAGTNAAR 42
|||:||||| |||:|||||:||||:||||:
Db 1271 GCCCAGCTTGGCCGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 37

US-11-169-041-32

; Sequence 32, Application US/11169041

; Publication No. US20060019284A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE

; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS

; FILE REFERENCE: 10001 NP

; CURRENT APPLICATION NUMBER: US/11/169,041

; PRIOR FILING DATE: 2005-06-28

; PRIOR APPLICATION NUMBER: 60/584,405

; PRIOR FILING DATE: 2004-06-30

; NUMBER OF SEQ ID NOS: 527

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 32

```
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-169-041-32

Query Match          90.4%; Score 30; DB 16; Length 2456;
Best Local Similarity 64.3%; Pred. No. 0.0063;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db 1271 GCCCAGCGCTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1312

RESULT 38
US-10-264-049-834
; Sequence 834, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patentn Ver. 3.1
; SEQ ID NO 834
; LENGTH: 2593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-049-834

Query Match          90.4%; Score 30; DB 7; Length 2593;
Best Local Similarity 64.3%; Pred. No. 0.0064;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db 1383 GCCCAGCGCTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1424

RESULT 39
US-10-198-846-10424
; Sequence 10424, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10424
; LENGTH: 4239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 4238, 4239
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10424

Query Match          90.4%; Score 30; DB 6; Length 4239;
Best Local Similarity 64.3%; Pred. No. 0.0074;
Matches 27; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db 1679 GCCCAGCGCTGGCGGAGCCCTTACAAAGACTATGAAGTAAAG 1720

RESULT 40
US-10-424-599-109942/c
; Sequence 109942, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 109942
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_70290C.1
US-10-424-599-109942

Query Match          67.5%; Score 22.4; DB 8; Length 365;
Best Local Similarity 59.0%; Pred. No. 12;
Matches 23; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 4 CAYGCGTGGCCNTCNCNTAYAAAGAYTAYGAGTNAAR 42
Db 181 CAGGAATGGCCCTCTCCCATAAAGATGATGACAAG 143

RESULT 41
US-09-925-065A-342206
; Sequence 342206, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 342206
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 4238, 4239
; OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-342206

Query Match          63.3%; Score 21; DB 4; Length 406;
Best Local Similarity 51.2%; Pred. No. 56;
```



```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 407850
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-407850
```

```
Query Match 63.3%; Score 21; DB 5; Length 552;
Best Local Similarity 51.2%; Pred. No. 61;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTCCGTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 308 CTCTTTCTCTCTCTCAATATGAAGATTATGAGGACAAA 268
```

```
RESULT 51
US-10-301-480-474416/c
; Sequence 474416, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 474416
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-474416
```

```
Query Match 63.3%; Score 21; DB 12; Length 559;
Best Local Similarity 51.2%; Pred. No. 62;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTCCGTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 315 CTCTTTCTCTCTCTCAATATGAAGATTATGAGGACAAA 275
```

```
RESULT 52
US-10-301-480-1087825/c
; Sequence 1087825, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
```

```
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1087825
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1087825
```

```
Query Match 63.3%; Score 21; DB 12; Length 559;
Best Local Similarity 51.2%; Pred. No. 62;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CNCAYGCGTCCGTCNCNTAYAAAGAYTAYGARGTNAAR 42
Db 315 CTCTTTCTCTCTCTCAATATGAAGATTATGAGGACAAA 275
```

```
RESULT 53
US-09-925-065A-42235/c
; Sequence 42235, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42235
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-42235
```

```
Query Match 63.3%; Score 21; DB 4; Length 619;
Best Local Similarity 55.3%; Pred. No. 63;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GNCAYGCGTCCGTCNCNTAYAAAGAYTAYGARGT 38
Db 39 GCCATTTCGACATCTCCACACAAAGTCTACGGACT 2
```

```
RESULT 54
US-09-925-065A-42235/c
; Sequence 42235, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
```

```
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42235
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-42235

Query Match          63.3%; Score 21; DB 5; Length 619;
Best Local Similarity 55.3%; Pred. No. 63;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAVCNTGCCNTCCNTAYAARGAYTAYGARGT 38
   |||:|||||:|||||:|||||:|||||:|||||:
Db 39 GCCATTCTGGACATCTCCACAAAGTCTACGGACT 2

RESULT 55
US-10-301-480-143473/c
; Sequence 143473, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143473
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-143473

Query Match          63.3%; Score 21; DB 12; Length 619;
Best Local Similarity 55.3%; Pred. No. 63;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAVCNTGCCNTCCNTAYAARGAYTAYGARGT 38
   |||:|||||:|||||:|||||:|||||:|||||:
Db 39 GCCATTCTGGACATCTCCACAAAGTCTACGGACT 2

RESULT 56
US-10-301-480-756882/c
; Sequence 756882, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
```

```
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 756882
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-756882

Query Match          63.3%; Score 21; DB 12; Length 619;
Best Local Similarity 55.3%; Pred. No. 63;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 GCNCAVCNTGCCNTCCNTAYAARGAYTAYGARGT 38
   |||:|||||:|||||:|||||:|||||:|||||:
Db 39 GCCATTCTGGACATCTCCACAAAGTCTACGGACT 2

RESULT 57
US-10-027-632-253967
; Sequence 253967, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253967
; LENGTH: 686
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-253967

Query Match          63.3%; Score 21; DB 6; Length 686;
Best Local Similarity 51.2%; Pred. No. 65;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 2 CMCAYCNTGCCNTCCNTAYAARGAYTAYGARGTNAAR 42
   |||:|||||:|||||:|||||:|||||:|||||:
Db 235 CTCTCTCAGCCTTCCCAATGTAAAGACGATGAGGTAAA 275

RESULT 58
US-10-027-632-253967
; Sequence 253967, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```



```

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63800
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-63800

```

Query Match 63.3%; Score 21; DB 5; Length 698;
Best Local Similarity 51.2%;
Pred. No. 66;
Matches 21; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CNCAYGCNTGGCCNCTCNCNTAYAAARGAYTAYGARGTNAAR 42
| | : | | | | | : | | : | | : | | :
D6 235 CTCTTCCTCAGCCTTCCC AATGTAAAGACGATGAGGTTAAA 275

RESULT 63

US-10-301-480-165038
; Sequence 165038, Application US/10301480
; Publication No. US20060057564A1

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome

```

```

; FILE REFERENCE: 108827, 137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1236818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165018
; LENGTH: 698

```

; LENGTH: 958
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-301-480-165038

Query Match	63.3%	Score 21;	DB 12;	Length 698;
Best Local Similarity	51.2%;	Pred. No. 66;		
Matches 21;	Conservative	7;	Mismatches 13;	Indels 0;
Gaps	0;			

Qy 2 CNcAYGcNTGGCCNtNCcCNTAYAAAGAYTAYGARGtNAAR 42
| | : | | | | | : | | : | | : | | :
Db 235 CTCTTCCTCAGCCTTCCCcAATGTAAAGACgATGAGGTATAA 275

RESULT 64

US-10-301-480-778447
; Sequence 778447, Application US/10301480
; Publication No. US20060057564A1

```
;
;
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
```

```

; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 778447
;

```

; LENGTH: 698
 ; TYPE: DNA
 ; ORGANISM: H
 US-10-301-480-7

Query Match 63.3%; Score 21; DB 12; Length 698;
Best Local Similarity 51.2%; Pred. No. 66;
Matches 21; Conservative 7; Mismatches 13; Indels

Qy 2 CNCAYGCNTGGCCNCTCNCNTATYAAARGAYTAYGAGGTNAAR 42
| | : | | | | | | | | : | | : | | : | | :
Db 235 CTCTTCCTCAGCCCTTCCCAATGTAAAGACGATGAGGTAAA 275

RESULT 65

US-10-767-701-19896/c
; Sequence 19896, Application US/10767701
; Publication No. US20040172684A1

;
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

1. TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
2. TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

```

; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128

```

```

; SEQ ID NO 198
; LENGTH: 710

```

```

;
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:

```

```

; NAME/KEY: unsure
; LOCATION: (1)..(710)
; OTHER INFORMATION: unsure at all n locations
;

```

FEATURE:
OTHER INFORMATION: Clone ID: LIB5049-011-R1-XP1-A5
US-10-767-701-19896

Query Match 63.3%; Score 21; DB 8; Length 710;
Best Local Similarity 58.3%; Pred. NO. 66;
Matches 21; Conservative 5; Mismatches 10; Indels

Qy 2 CNCAYGCNTGGCCNCTCNCNTAYAAAGGATYARG 37
| | | | | | | | | | : | | | | |
Db 131 CGCGCGGTAGCCGTGCGCCGACAAAGGACGACGACG 96
| | | | | | | | | | : | | | | |

RESULT 66

US-10-932-182A-82765
; Sequence 82765, Application US/10932182A
; Publication No. US20060046253A1

;
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA

; APPLICANT: KODAMA, YUKIKO
 ; APPLICANT: FUJIMURA, TOMOKO
 ; APPLICANT: ASHIKARI, TOSHIHIKO

; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 ; FILE REFERENCE: 030685-043
 ; CURRENT APPLICATION NUMBER: US/10/932,182A
 ; CURRENT FILING DATE: 2004-09-02

```

; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82765

```

; LENGTH: 1101
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces pastorianus
 US-10-932-182A-82765

Query Match	63.3%	Score 21;	DB 11;	Length 1101;
Best Local Similarity	55.3%	Pred. No. 75;		


```
RESULT 71
US-10-451-467A-363
; Sequence 363, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKWANS, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; TITLE OF INVENTION: YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 363
; LENGTH: 2780
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-363

Query Match 61.4%; Score 20.4; DB 8; Length 2780;
Best Local Similarity 56.8%; Pred. No. 1.8e+02;
Matches 21; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 6 YGCGTGGCCNTCCNTAAGAYTAYGARGTNAAR 42
DB 2716 CGCCTGGCCATCACCATACATTGGTTAGGAGGAAAA 2752

RESULT 72
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; CURRENT FILING DATE: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 61.4%; Score 20.4; DB 3; Length 3309400;
```

```
Best Local Similarity 52.5%; Pred. No. 1.2e+03;
Matches 21; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 GCNCAYGCGTGGCCNTCCNTAYAAARGAYTAYGARGTNA 40
DB 552828 GCGCAGCGCTTGACCACGCGCTTGTGAAGACGACCAAGTGA 552789

RESULT 73
US-10-719-993-31292/c
; Sequence 31292, Application US/10719993
; Publication No. US2004040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31292
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-31292

Query Match 60.8%; Score 20.2; DB 9; Length 201;
Best Local Similarity 56.4%; Pred. No. 1.1e+02;
Matches 22; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 2 CMCAYGCGTGGCCNTCCNTAYAAARGAYTAYGARGTNA 40
DB 88 CACACGCTTCCCATCTCCTTTCAATGAATGTGTGTCA 50

RESULT 74
US-10-282-122A-25373/c
; Sequence 25373, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 25373
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Legionella pneumophila
US-10-282-122A-25373

Query Match      60.8%; Score 20.2; DB 8; Length 1146;
Best Local Similarity 53.7%; Pred. No. 1.8e+02;
Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GCNCAYCGTCGTCNCNTCCNTAAYAAAGAYTAYGARGTNA 41
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1081 GCCATAAGTTGCGTCTGCTATATAATAGACGAAGCAA 1041

RESULT 75
US-10-719-993-6844/c
; Sequence 6844, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: CL001496
; CURRENT FILING DATE: US/10/719,993
; NUMBER OF SEQ ID NOS: 59342
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 6844
; LENGTH: 775062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(775062)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6844

Query Match      60.8%; Score 20.2; DB 9; Length 775062;
Best Local Similarity 56.4%; Pred. No. 1.1e+03;
Matches 22; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CNCAYCGTCGTCNCNTCCNTAAYAAAGAYTAYGARGTNA 40
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 252253 CACACGCTTTCCCATCTCTTTCATGATGATGTTGGTCA 252215

RESULT 76
US-10-461-862-166/c
; Sequence 166, Application US/10461862
; Publication No. US20050090434A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 52945201800
; CURRENT APPLICATION NUMBER: US/10/461,862
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 783062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)...(783062)
; OTHER INFORMATION: n = A,T,C or G
US-10-461-862-166

Query Match      60.8%; Score 20.2; DB 10; Length 783062;
Best Local Similarity 56.4%; Pred. No. 1.1e+03;
Matches 22; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CNCAYCGTCGTCNCNTCCNTAAYAAAGAYTAYGARGTNA 40
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 256252 CACACGCTTTCCCATCTCTTTCATGATGATGTTGGTCA 256214

RESULT 77
US-10-674-124A-11800/c
; Sequence 11800, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: GRIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 11800
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AC004593.1_19975
; FEATURE:
; OTHER INFORMATION: Located on chromosome 7
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 30076167
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 8009
US-10-674-124A-11800

Query Match      60.2%; Score 20; DB 9; Length 282;
Best Local Similarity 54.1%; Pred. No. 1.5e+02;
Matches 20; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 4 CAYCGTCGTCNCNTCCNTAAYAAAGAYTAYGARGTNA 40
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 272 CATGGCAGGACATCACGGTACAAAGGCTATGAGGCAA 236

RESULT 78
US-10-437-963-66738/c
; Sequence 66738, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
```

```
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 66738
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67661C.1
US-10-437-963-66738

Query Match      60.2%; Score 20; DB 8; Length 381;
Best Local Similarity 54.1%; Pred. No. 1.8e+02;
Matches 20; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY      6 YGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db      105 TGCATGACCTCAACAGATATAGTACGAGTAGAA 69

RESULT 79
US-09-925-065A-219816/c
; Sequence 219816, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219816
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-219816

Query Match      60.2%; Score 20; DB 4; Length 560;
Best Local Similarity 51.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      4 CAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
      ||:|||:|||:|||:|||:|||:|||:|||:|||:
Db      270 CATGCTGGCCACCACCAATATCAACATTTTAAATGAAA 232

RESULT 80
US-09-925-065A-219817/c
; Sequence 219817, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
```

```
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219817
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-219817

Query Match      60.2%; Score 20; DB 4; Length 560;
Best Local Similarity 51.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      4 CAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
      ||:|||:|||:|||:|||:|||:|||:|||:|||:
Db      270 CATGCTGGCCACCACCAATATCAACATTTTAAATGAAA 232

RESULT 81
US-09-925-065A-219818/c
; Sequence 219818, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219818
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-219818

Query Match      60.2%; Score 20; DB 4; Length 560;
Best Local Similarity 51.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY      4 CAYGCGTGGCCNTCCNTAYAAAGAYTAYGARGTNAAR 42
      ||:|||:|||:|||:|||:|||:|||:|||:|||:
Db      270 CATGCTGGCCACCACCAATATCAACATTTTAAATGAAA 232

RESULT 82
US-09-925-065A-219816/c
; Sequence 219816, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```

; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 219816
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-219816

```

```
Query Match      60.2%; Score 20; DB 5; Length 560;
Best Local Similarity 51.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
```

Qy	4	CAYGCNTGGCCNTCNCNTAYAARGAYTAYGARGTNAAR	42
		: : : : : :	
D_b	270	CATGCCTGGCCCA CCA CAATATCAACATTTTAAATGAAG	232

RESULT 83

US-09-925-065A-219817/c
; Sequence 219817, Application US/09925065A
; Publication No. US20050228172A9

```

1  ; ORGANISM: Homo sapiens
2  ; ORGANISM: Homo sapiens
3  ; US-09-925-065A-219817
4  ;
5  ; TYPE: DNA
6  ; LENGTH: 560
7  ; SEQ ID NO 219817
8  ; SOFTWARE: FastSeq for Windows Version 4.0
9  ; NUMBER OF SEQ ID NOS: 957086
10 ;
11 ; PRIOR FILING DATE: 2001-05-09
12 ; PRIOR APPLICATION NUMBER: US 60/289,846
13 ;
14 ; PRIOR FILING DATE: 2001-01-16
15 ; PRIOR APPLICATION NUMBER: US 60/261,766
16 ;
17 ; PRIOR FILING DATE: 2000-11-30
18 ; PRIOR APPLICATION NUMBER: US 60/250,092
19 ;
20 ; PRIOR FILING DATE: 2000-11-20
21 ; PRIOR APPLICATION NUMBER: US 60/252,147
22 ;
23 ; PRIOR FILING DATE: 2000-10-24
24 ; PRIOR APPLICATION NUMBER: US 60/243,096
25 ;
26 ; PRIOR FILING DATE: 2001-08-08
27 ; CURRENT FILING DATE: 2001-08-08
28 ;
29 ; CURRENT APPLICATION NUMBER: US/09/925,065A
30 ;
31 ; FILE REFERENCE: 108827.135
32 ;
33 ; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
34 ;
35 ; APPLICANT: Wang, David G.
36 ;
37 ; ORGANIZATION:
38 ;
39 ;
40 ;
41 ;
42 ;
43 ;
44 ;
45 ;
46 ;
47 ;
48 ;
49 ;
50 ;
51 ;
52 ;
53 ;
54 ;
55 ;
56 ;
57 ;
58 ;
59 ;
60 ;
61 ;
62 ;
63 ;
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;

```

```
Query Match      60.2%; Score 20; DB 5; Length 560;
Best Local Similarity 51.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
```

Qy 4 CAYGNTGGCCNCTNCNCNTAYAARGAYTAYGARGTNAAR 42
||:||||||| | ||:|.:.| : ||:
Dp 270 CATGCCTGGCCACCACAATATCAACATTCTTTTAATGAAG 232

RESIT.T 84

RESULI 84
US-09-925-065A-219818/c
; Sequence 219818, Application US/09925065A

```

; Publication No. US20050228172A9
;
; GENERAL INFORMATION:
;
; APPLICANT: Wang, David C.
;
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
;
; FILE REFERENCE: 108827.135
;
; CURRENT APPLICATION NUMBER: US/09/925,065A
;
; CURRENT FILING DATE: 2001-08-08
;
; PRIOR APPLICATION NUMBER: US 60/243,096
;
; PRIOR FILING DATE: 2000-10-24
;
; PRIOR APPLICATION NUMBER: US 60/252,147
;
; PRIOR FILING DATE: 2000-11-20
;
; PRIOR APPLICATION NUMBER: US 60/250,092
;
; PRIOR FILING DATE: 2000-11-30
;
; PRIOR APPLICATION NUMBER: US 60/261,766
;
; PRIOR FILING DATE: 2001-01-16
;
; PRIOR APPLICATION NUMBER: US 60/289,846
;
; PRIOR FILING DATE: 2001-05-09
;
; NUMBER OF SEQ ID NOS: 957086
;
; SOFTWARE: Fast-SEQ for Windows Version 4.0
;
; SEQ ID NO 219818
;
; LENGTH: 560
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-09-925-065A-219818

```

Query Match 60.2% Score 20; DB 5; Length 560;
Best Local Similarity 51.3% Pred. No. 1.8e+02;
Matches 20; Conservative 12; Indels 0; Gaps 0;

Qy	4	CAYGCNTGGCCN	TNC	CNC	N	TAY	AARG	AYTAYG	GARGTNAAR	42
		:		:	:	:	:	:	:	
Db	270	CATGCCTGGCC	ACC	CA	CA	TAT	CA	ACATTTT	TAATGA	232

RESULT 85

RESUB 83
US-10-301-480-305477/c
; Sequence 305477, Application US/10301480
; Publication No. US20060057564A1

```

? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
? TITLE OF INVENTION: in the Human Genome
? FILE REFERENCE: 108827.137
? CURRENT APPLICATION NUMBER: US/10/301,480
? CURRENT FILING DATE: 2002-11-21
? PRIOR APPLICATION NUMBER: US 10/215,598
? PRIOR FILING DATE: 2002-08-09
? PRIOR APPLICATION NUMBER: US 60/311,695
? PRIOR FILING DATE: 2001-08-10
? NUMBER OF SEQ ID NOS: 1226818
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 305477
? LENGTH: 575
? TYPE: DNA
? ORGANISM: Homo sapien
? US-10-301-480-305477

```

Query Match	60.2%	Score 20;	DB 12;	Length 575;
Best Local Similarity	51.3%	Pred. No.	1.8e+02;	
Matches 20: Conservative		7; Mismatches 12;	Indels 0;	Gaps 0;

QY 4 CAYGNTGGCCNTCNCNTAYAARGAYTAYGARGTNAAR 42
||:||||||| ||:||:||:||:||:||:||:
nb 270 CATGCCTGGCCACCACAATATCAACATTTTTAAATGAAA 232

DEC 17 1966

RESULT 86
 US-10-301-480-305478/c
 ; Sequence 305478, Application US/10301480
 ; Publication No. US2006005756A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.

Best Local Similarity 51.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

Qy

4 CAYGCNTGGCCNCTNCNCNTAYAARGAYTAYGARGTNAAR 42
||:||| ||||| | | ||: | : | : | : | :

Db

270 CATGCCTGGCCACCAATATCAAATTATTTTAATGAATA 232

```

RESULT 91
US-10-312-273-340/c
, Sequence 340, Application US/10312273
, Publication No. US20040005667A1
, GENERAL INFORMATION:
, APPLICANT: CHIRON SPA
, TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
, FILE REFERENCE: P025035WO
, CURRENT APPLICATION NUMBER: US/10/312,273
, CURRENT FILING DATE: 2002-12-20
, PRIOR APPLICATION NUMBER: 0016363.4
, PRIOR FILING DATE: 2000-07-03
, PRIOR APPLICATION NUMBER: 0017047.2
, PRIOR FILING DATE: 2000-07-11
, PRIOR APPLICATION NUMBER: 0017983.8
, PRIOR FILING DATE: 2000-07-21
, PRIOR APPLICATION NUMBER: 0019368.0
, PRIOR FILING DATE: 2000-08-07
, PRIOR APPLICATION NUMBER: 0020440.4
, PRIOR FILING DATE: 2000-08-18
, PRIOR APPLICATION NUMBER: 0022583.9
, PRIOR FILING DATE: 2000-09-14
, PRIOR APPLICATION NUMBER: 0027549.5
, PRIOR FILING DATE: 2000-11-10
, PRIOR APPLICATION NUMBER: 0031706.5
, PRIOR FILING DATE: 2000-12-22
, NUMBER OF SEQ ID NOS: 664
, SOFTWARE: SeqWin99, version 1.02
, SEQ ID NO 340
, LENGTH: 669
, TYPE: DNA
, ORGANISM: Chlamydia pneumoniae
US-10-312-273-340

```

```

RESULT 92
US-09-764-869-1560/c
; Sequence 1560, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1560
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1560

```

Qy 4 CAYGNTGGCCNTCCNTAYAAAGAYTAYGA 35
||| ||| ||| ||| : ||| : |||
Db 4936 CATGCTGGCCGTCATATTCCTAATGATTATGA 4905

```

RESULT 93
US-10-091-504-1560/c
; Sequence 1560, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See file Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1560
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1560

```

Query Match	60.2%	Score 20;	DB 6;
Best Local Similarity	62.5%	Pred. No. 3.4e+02	
Matches	20;	Conservative	4; Mismatches
QY	4	CAYGNTGGCCNTCNCNTA	YAARGAYTAYGA 35
Db	4936	CATGCTGGCCGTCATATCT	TAATGATTAATGA 4905

```

RESULT 94
US-10-227-577-1560/c
; Sequence 1560, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1560
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-227-577-1560

```

```
Query Match      60.2%; Score 20; DB 7; Length 5161;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 20; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      4 CAYGCTGCGCCNTCCNTAATGAGATATGA 35
      ||:|||||:|||||:|||||:|||||:|||||:
Db      4936 CATGCTGCGCCGTCATATTCATATGATTA 4905

RESULT 95
US-10-289-762-1
; Sequence 1, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (165001)..(180000)
```

```
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
```

```
Query Match      60.2%; Score 20; DB 7; Length 1230025;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 20; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY      2  CNCAYGNTGCCNTCCNTAARGAYTAYGAR 36
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      835280  CTCATGCATAGGCTTTACCATAGAAAGATAACGAG 835314
```

```
RESULT 96
US-09-925-065A-519993/c
; Sequence 519993, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 519993
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-519993
```

```
Query Match      59.6%; Score 19.8; DB 4; Length 570;
Best Local Similarity 55.3%; Pred. No. 2.2e+02;
Matches 21; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY      5  AYCNTGCCNTCCNTAARGAYTAYGARGTNAAR 42
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      454  ATCGATGGCATTAAAGGCTAAAGATTAGCACTTAAG 417
```

```
RESULT 97
US-09-925-065A-519993/c
; Sequence 519993, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 519993
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-519993
```

```
Query Match      59.6%; Score 19.8; DB 5; Length 570;
Best Local Similarity 55.3%; Pred. No. 2.2e+02;
Matches 21; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
```

